



UNIT 4 PRACTICAL WORK IN SCIENCE

STUDY GUIDE	3	4 THE STATISTICAL APPROACH: STANDARD DEVIATION	25
1 INTRODUCTION	3	Summary of Section 4	27
2 SCIENTIFIC INVESTIGATION	4	5 WRITING REPORTS OF PRACTICAL WORK (AV SEQUENCE)	27
2.1 Experiments of observation and experiments of measurement	4	OBJECTIVES FOR UNIT 4	35
2.2 Designing experiments	4	ITQ ANSWERS AND COMMENTS	35
2.3 Planning your practical work for the Course	6	SAQ ANSWERS AND COMMENTS	37
2.4 Keeping records	7	ACKNOWLEDGEMENT	38
Summary of Section 2	8	INDEX FOR UNIT 4	38
3 HANDLING EXPERIMENTAL DATA	9		
3.1 Making use of graphs	9		
3.2 Identifying experimental uncertainties	16		
3.3 Accuracy and precision in experimental work	17		
3.4 Estimating experimental uncertainties	19		
3.5 Quoting experimental results: uncertainties and significant figures	21		
3.6 Graphical representation of uncertainties	22		
Summary of Section 3	23		

STUDY GUIDE

This Unit is somewhat unusual in that it introduces very few results or laws. Rather, its purpose is to make you think more deeply about how you, as a scientist, should carry out practical work, record your observations and interpret any data you collect. You will find that some of the material presented in this Unit is already familiar—a number of the ideas featured in Units 2 and 3 and in the preparatory package. Some of these ideas were not made much of where they first appeared, because that might have obscured the main arguments of the early Units; others may only make sense to you after you have had more chance to practise the skills involved. Unit 4 is intended to give you the opportunity to draw all the threads together and to consolidate some of the skills that, as a science student, you will need to exercise frequently. Parts of the Unit should also prove useful for revision and reference throughout your study of the Course.

This Unit has intentionally been kept short, to allow you time for the other major component of this week's work—the completion of TMA 01. The AV sequence (Section 5) is designed to help you with this assignment.

The TV programme 'Practically speaking' is associated with this Unit, and it may be watched with equal profit at any stage of your progress through the text.

I INTRODUCTION

Observation and the collection of data are essential to most progress in science. In this Unit we shall address fundamental questions about how data should be collected, reported and analysed.

We begin, in Section 2, by looking at various types of practical work; we shall consider what features of scientific method are common to work in all disciplines, and what special problems are associated with particular branches of scientific investigation. A biologist probably won't approach experiments in quite the same way as a physicist, but they will, nevertheless, share a similar careful and critical attitude to practical work. By the end of your study of this Course, you too should have adopted this attitude! The TV programme 'Practically speaking' is particularly relevant to Section 2.

Section 3 is devoted to the analysis of quantitative data, with discussion of such questions as 'How do you collect and interpret data?', 'What confidence limits do you place on your final results?' and 'How do you present those results to other people?'. Again, the TV programme shows some pertinent examples, drawn from a variety of scientific disciplines.

Section 4 looks briefly at a more sophisticated method of data analysis. If you are very short of time for the week's work, this is the Section to skip, since it is not essential to your progress through later Units; however, it does bring out one vital point of good practice in experimental work, so you should at least read the summary carefully.

Finally, Section 5 consists of an AV sequence that will help you put all these ideas into practice when you write the report required for TMA 01.

2 SCIENTIFIC INVESTIGATION

2.1 EXPERIMENTS OF OBSERVATION AND EXPERIMENTS OF MEASUREMENT

Practical work in science comes in many forms. Many of the examples you met in Units 2 and 3 involved experiments in which the end product was a *quantitative* (i.e. numerical) result—such as an estimate of the magnitude of the acceleration due to gravity. On the other hand, you have also met practical work of a more *qualitative* nature—for example, your observations of the behaviour of the Foucault pendulum (Unit 1). Both qualitative and quantitative work have their place in scientific investigation, and you will meet many examples of each type during your study of this Course.

Careful observation is the basis of most scientists' craft. An enormous amount of knowledge can be amassed without making a single measurement. An Earth scientist can reconstruct the geological history of an area over a period of millions of years, by noting what types of rock are present and how they overlie or intersect one another. The structure of the rocks can reveal how they have been shifted as a result of large-scale movements in the Earth's crust, and a geologist can also 'read' the landscape to work out how the rocks have subsequently been changed by the effects of wind, water or ice. Many chemical investigations are also based on observation, rather than quantitative measurement. In Units 13–14 and at Summer School you will perform a number of experiments of this type, systematically recording what happens when various substances are allowed to react (e.g. a colour change, deposition of a solid from a solution, or emission of a gas). It is on the accumulation of vast quantities of such evidence that much of our modern understanding of chemistry rests.

Observation, coupled with accurate reporting and interpretation of observations, is vitally important, but precise measurements are also crucial, and much of this Unit will be taken up with matters relating to quantitative experiments. Even so, it is not enough simply to acquire data. One has to be able to analyse the results and assess their reliability. The handling of experimental data is the main topic of Section 3.

However, the process of investigation begins long before the first measurement is made. In the remainder of this Section, therefore, we shall consider first the initial planning phase for practical work, and then the best way to document the work as it proceeds.

2.2 DESIGNING EXPERIMENTS

To begin with, it is important to be quite sure of what it is that is actually being observed or measured—the act of making the observation may itself alter what is being observed. This is particularly obvious in some biological investigations: as soon as an attempt is made to study the behaviour of an animal under controlled conditions, the very fact that it is being observed may affect its behaviour. Similar problems can arise in all branches of science. If the pressure in a tyre is being measured, the attachment of the pressure gauge releases some of the air from the tyre and so alters the pressure being measured. If the diameter of a rubber ball is being measured with a micrometer, the pressure exerted by the jaws of the instrument may squash the ball, so that a false value for the diameter will be obtained (Figure 1). It is always important to watch out for, and if possible reduce, problems of this nature.

Next, there is the question of how best to choose the samples that are to be investigated. No matter how carefully we make our observations, if the sample being tested is unrepresentative, the effort is wasted.

For example, a geologist examining fossils from a particular rock formation might decide to examine only complete shells. From one point of view, this seems perfectly reasonable, because whole fossils can reveal more than

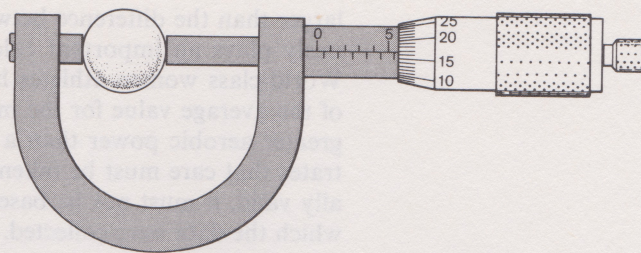


FIGURE 1 Using a micrometer to measure the diameter of a small rubber ball. The kind of scale engraved on this gauge—called a Vernier scale—allows the diameter to be measured to within 0.001 cm. You will get experience of reading Vernier scales at Summer School.

broken ones of the same type. However, the complete shells might not be representative of the fossil assemblage of that rock bed as a whole: they might simply be much more robust than the other animal and vegetable remains in the bed. Similarly, when making a geological map, one has to be careful not to overestimate the proportion of the harder rocks; because these are more resistant to erosion, they may form conspicuous solid outcrops and so can appear to be the most characteristic feature of the area.

A common problem in biology is illustrated by the following example. Suppose you wanted to monitor the population of shore birds on an estuary, and that you did this by counting the number of birds along a given stretch of the shore line. Clearly, if you made the count when the tide was in and the birds' feeding grounds were covered, you would not get the same result as if you made the count when the tide was out. This illustrates that the particular conditions under which a sample is taken may introduce a bias, which might invalidate any conclusions based on the sample.

Another difficulty often encountered in biological experiments is discussed in the TV programme 'Practically speaking'. This difficulty arises from the natural variability of living organisms. Imagine you wanted to measure the fitness of a group of people by seeing how they responded to a burst of physical exercise (Figure 2). One way of doing this would be to monitor each person's uptake of oxygen during exercise, which is a measure of the efficiency of their heart and lungs; the technical term for this is 'aerobic power'. You might expect the results to show a significant difference between the average aerobic power of men and that of women—and this expectation has, in fact, been borne out by experiment. Typically, values of maximum aerobic power are 25–30% lower in women than in men. However, the variation between individuals of the same sex is sometimes



FIGURE 2 The response of a randomly-selected group of people to a physical challenge.

larger than the difference between the averages for each sex. Training obviously plays an important role, so does age, and so does 'natural ability'. World-class women athletes have maximum aerobic powers well in excess of the average value for the male population; a typical 17-year-old girl has greater aerobic power than a typical 60-year-old man. This example illustrates that care must be taken to ensure that if a conclusion is to be generally valid, it must not be based on a result that is a function of the way in which the data were collected.

This kind of concern with sampling techniques and with the natural variability within and between samples can make the approaches of biologists and geologists rather different to those of physicists and chemists. As you have seen in Units 2 and 3, physicists usually expect the results of an experiment to be reproducible, and therefore they normally estimate the uncertainty of quantitative results from the uncertainties involved in making the individual measurements. We will return to the quantitative treatment of such uncertainties in Section 3.

2.3 PLANNING YOUR PRACTICAL WORK FOR THE COURSE

It's an old adage that time spent in reconnaissance is seldom wasted. It certainly applies to practical work in science: effort put in at the planning stage will pay handsome dividends later on. Also, the amount of time you have for carrying out experiments will probably be limited, and it's important to use it to your best advantage.

Before you start any piece of practical work, try to get a clear idea of what you are aiming to do in the experiment, and read any instructions thoroughly. You will then know which are the crucial measurements, and will be able to concentrate on making these as carefully as possible. You will also avoid wasting time by being unnecessarily precise about unimportant details. Suppose you have to measure out a particular volume of liquid. If the liquid were to be used in a medical context, where the volume was small and accuracy all-important, you might use a syringe; the one illustrated in Figure 3a could be used to measure a volume of 0.3 ml to within 0.01 ml. In a chemistry experiment, on the other hand, it might be sufficiently accurate to measure the volume in a graduated cylinder; using the one shown in Figure 3b you could measure out 16 ml to within 0.5 ml. In some cases, a rough visual estimate of volume may be adequate, and then you could simply use a beaker (Figure 3c).

Many decisions of this kind, and others that may occur to you while planning an experiment, can be made after doing a preliminary experiment. It

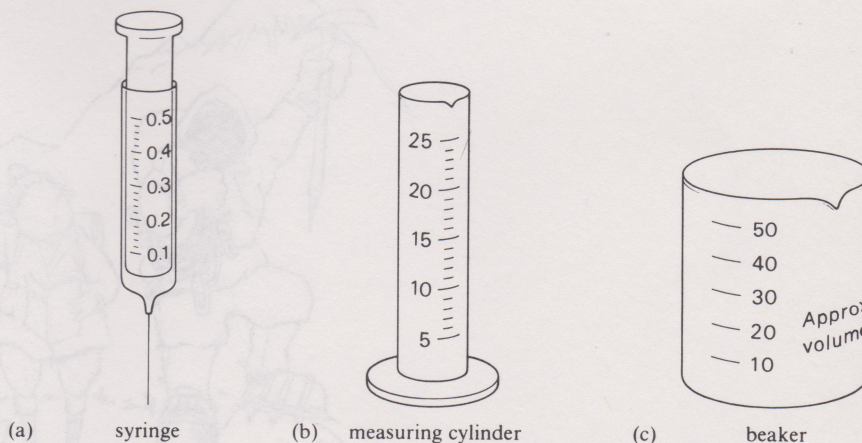


FIGURE 3 Various items of glassware that can be used to measure the volume of a liquid, *not* drawn to the same scales. (Glassware is usually graduated in ml or cm^3 , and to all intents and purposes these two sets of measures are equivalent: $1 \text{ ml} = 1 \text{ cm}^3$.)



FIGURE 4 This experiment, devised by a student to improve his measurement of the distance to the Moon, is described in the TV programme 'Practically speaking'. (Note that if you were actually using this method to measure the diameter of the clock face, it would be essential to position the ladder so that you were looking at the clock square on.)

often pays to have a quick rehearsal* before starting the experiment proper; it *always* pays to set up the apparatus to check that everything is functioning correctly. If you can have a brief trial run, you will develop a 'feel' for the experiment, and you will also have preliminary answers to questions such as 'What is the range over which measurements should be taken?', 'What is a suitable interval between measurements?' and 'From what are the largest uncertainties likely to arise, and can anything be done to minimize them?'.

It's rather like planning a holiday journey. The instructions for the experiment form your basic route-map, showing the essential steps required to get from A to B. If you follow the main road from A to B on the map, you'll get to your destination without difficulty. However, you should also aim to enjoy the countryside on the way. If you see a side turning or a possible short-cut, investigate it. Be flexible and let your natural curiosity lead you occasionally. If you can see a way of improving the experiment, try it out (Figure 4). Above all, look critically at what you are doing and keep an open mind. However, don't get so carried away that you forget the original aim of the experiment or the safety instructions!

2.4 KEEPING RECORDS

A scientist's field or lab notebooks (Figure 5) are highly valued possessions. They are literally irreplaceable. Not only do they often represent months or years of work, but also the information they contain is the product of a unique set of circumstances that can never be re-created in every particular.

As a science student, you too should treat your own observations and data with due respect. It should be a cardinal rule to record all your practical work in a bound notebook, and to write everything into it *as the experiment proceeds*. It is not good practice to use loose sheets of paper—they can (and

7/5/72	<u>MIGRATION COUNT</u>	18	R. b. Merganser 7/4/11
	Dungeness TR 0916	68	C/A Tern
	Wind SSW force 3		6/7/10/16/2/11/16
	2/8 cloud. Mild.	148	Sandwich Tern
	6:30-11:00 AM		12/17/30/31/32/12/14
③	Kittiwake	18	Little Tern 2/3/1/12
65	Gannet 19/23/7/7/7/2	4	Black Tern 2/1/1
4	Whimbrel	8	B-T Diver
662	Common Scoter	1	R-T Diver
	174/179/57/54/80/8/110	8	Bar.T. Godwit 2/6
51	Velvet Scoter	37	Oystercatcher 6/31
	2/15/4/3/1/6	1	Razorbill
9	Arctic Skua	5	Little Gull 1/2/2
	172/1/3/1/1	3	Skua sp.
10	Pom. Skua 3/5/2	4	Diver sp.
4	Bonxie 1/2/1		

FIGURE 5 A page from a field notebook. (The figures after each species show the numbers of birds seen, group by group; the ringed figure is the total number for that species.)

* Of course if you are using chemicals in your experiment, you have a limited amount in the Experiment Kit, and if you are growing plants, it takes a long time, so a full rehearsal may not be feasible.



FIGURE 6 It is unwise to take field notes on loose sheets of paper.

almost certainly will!) get lost (Figure 6). Punched sheets and a ring binder can be made to serve as a last resort, but they are far from ideal. Rough paper or the backs of envelopes, however, are an anathema.

It is a good habit to note the date and time at the start of each set of observations; for field-work a grid reference should also be included, as shown in Figure 5. Remember, too, that the unexpected, inexplicable or otherwise 'odd' observation may eventually turn out to be of crucial significance, so don't ignore it—make a note of it for future reference.

You shouldn't be afraid to spread your notes over several pages. If your observations are set out clearly and are well spaced, the time you save when you need to refer to them will compensate handsomely for the paper you use. It is never a waste of paper to jot down notes about the progress of your experiment or your observations, or to set out your measurements properly. It is remarkable how a few cryptic comments or numbers, which seemed completely self-explanatory when they were written, become utterly meaningless after a very short time.

Measurements are most easily recorded in the form of a Table. Remember to head each column!

ITQ 1 The mean distances of three planets from the Sun, and their orbital periods, are: Mercury, 0.058×10^{12} m, and 0.24 Earth year; Venus, 0.108×10^{12} m, and 0.62 Earth year; Earth, 0.150×10^{12} m, and 1 Earth year. Set these data in a headed Table.

So, the most important characteristics of a lab notebook are that it should constitute as complete a record as possible of the experimental work and that it should be clear (to your eager biographer in the next century as well as to you personally in the short and long term!).

In later Sections and in the TV programme, you will see lab records for a variety of experiments. Meanwhile, you might like to look at your own notes for the determination of the distance to the Moon (and for the preparatory cross experiment, if you did that as well). Do you need to make any adjustments to your working practices?

SUMMARY OF SECTION 2

Experiments must be carefully designed if the conclusions based upon them are to have general validity. A 'trial run' may be helpful in highlighting problems or suggesting improvements. You should keep a permanent record of all your practical work in your Notebook.

3 HANDLING EXPERIMENTAL DATA

In this Section we shall be concerned mainly with the type of experiment that we expect to give quantitative, reproducible results. We've already given some thought to obtaining and recording the data, but this is only the first (and in some ways the easiest) stage of the process. The next steps require decisions about the following:

Presenting the data Often this can best be done graphically, so in Section 3.1 there are a few preliminary reminders about graphs.

Estimating uncertainties in the measurements In Section 3.2 we shall review the ways in which experimental uncertainties arise, and in Section 3.3 we shall look at techniques for quantifying them. We shall discuss how to present uncertainties in Sections 3.4, 3.5 and 3.6, while in Section 4 we shall look at more sophisticated techniques for analysing uncertainties.

3.1 MAKING USE OF GRAPHS

Imagine that you are doing some research into the behaviour of copper wires under load. You begin by trying to find out how a particular sample stretches when a load is applied to the end of it, so you set up an experiment in which you hang a series of weights on the end of the wire, and measure the increase in length resulting from each successive increase in mass. Figure 7 shows what the relevant page of your laboratory notebook might look like. (There is no need at this stage to worry about the practical details, but you may be interested to know that a travelling microscope is a device for measuring small changes in length.)

10/12/87

STRETCHING OF COPPER WIRES

Gauge 000 wire from reel (A) (non-enamelled).
 Very small weight attached to straighten
 out kinks in the wire before 'unstretched'
 length measured. Additional weights
 only are recorded.

For 'unstretched' wire
 travelling microscope reading = 1.2 mm

mass/kg	reading/mm	extension/mm
5.0	1.4	0.2
10.0	1.7	0.5
15.0	2.0	0.8
20.0	2.2	1.0
22.5	2.4	1.5
25.0	2.5	1.3
27.5	2.6	1.4
30.0	2.7	1.5
32.5	2.9	1.7
35.0	3.0	1.8
37.5	3.1	1.9
40.0	3.2	2.0
42.5	3.5	2.3
45.0	3.7	2.5
47.5	4.0	2.8
50.0	4.4	3.2

Manufacturer's tolerance on weights
 = 0.1% (check with balance later).
 Travelling microscope readings ± 0.05 mm

FIGURE 7 A record of an experiment.

There are a number of points worth noting about the record shown in Figure 7:

- (i) It does not consist merely of data; there are also brief reminders about the procedure adopted.

INDEPENDENT VARIABLE

DEPENDENT VARIABLE

INTERPOLATION

(ii) The initial microscope reading for the unstretched wire has been noted, as have all the microscope measurements for the wire under load. The extensions have been calculated subsequently. It is wise always to record the readings, including the zero reading, exactly as they are made. The differences can then be entered in a separate column, which can be checked for possible mistakes. Nobody is completely immune from the odd arithmetic slip!

(iii) Although uncertainties have not been quoted for every reading, notes have at least been made on the precision expected for the measurements of mass and length. These could be used at a later stage to estimate overall confidence limits. (We will return to the estimation of uncertainties in Section 3.4.)

If you obtained the results shown in Figure 7, what conclusions would you draw? In fact, it is not at all easy to extract information from tabulated data, but the picture becomes much clearer when the results are plotted on a graph, as in Figure 8. By now, some of the essential features of a well-

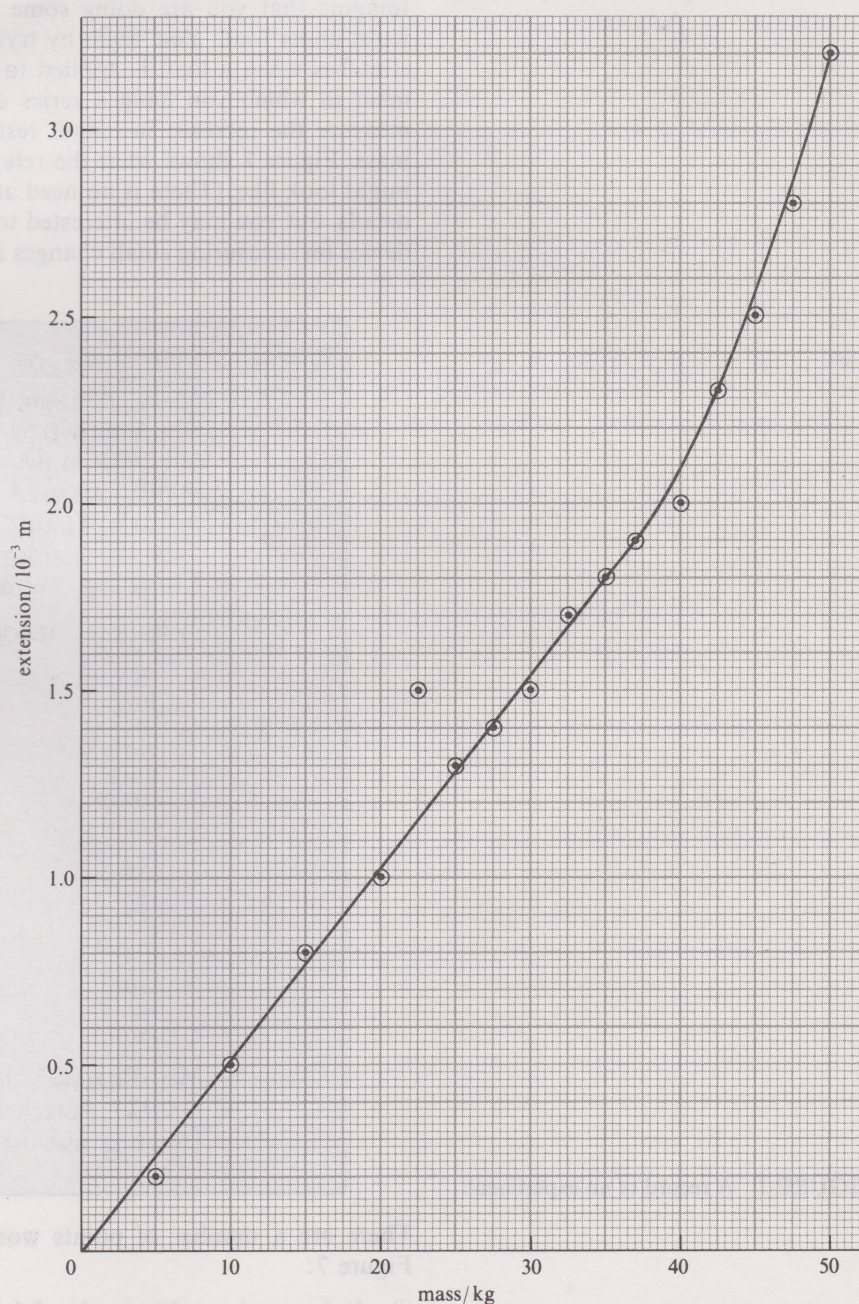


FIGURE 8 A graph to show the extension of a copper wire under load.

drawn graph should be quite familiar to you. In particular, note the following:

(i) Both axes on Figure 8 are labelled with the quantity divided by the appropriate units (and a power of ten where necessary), so that what are actually plotted are simply dimensionless numbers.

(ii) Figure 8 follows the normal convention: the **independent variable** (i.e. the quantity that is under the control of the experimenter—in this case the mass) is plotted along the horizontal axis, and the **dependent variable** (i.e. the resulting effect) is plotted along the vertical axis.

(iii) The points on Figure 8 are plotted in such a way that they show up clearly. Tiny dots made with a hard pencil or fine pen are not easy to see; huge blobs result in a loss of precision. Crosses ($\times$ or $+$), or dots with circles round them ($\odot$) are the best solutions.

Seeing results presented on a graph makes them much easier to interpret. Just by glancing at Figure 8, we can answer a number of questions about the extension of this particular wire under load:

- ☐ How is the extension related to the magnitude of the load?
- For masses up to about 35 kg the graph is a straight line, from which it follows that the extension is proportional to the mass; that is, doubling the mass produces twice the extension. For masses greater than about 35 kg, the graph is no longer a straight line, so the extension is no longer proportional to the mass; regular increases in mass produce successively greater increases in extension.

- ☐ Are the results physically reasonable?

■ A sensible thing to do in this context is to extend the graph to the left of the first reading; this has been done in Figure 8 and it is clear that the extended straight line passes through the origin. This is exactly what you would expect—with no load there should be no extension. This kind of procedure (which you will remember from Unit 2 is called *extrapolation*) can often act as a very useful check.

- ☐ Should any of the data points be re-checked?

■ The point plotted as mass = 22.5 kg, extension = 1.5 mm, appears to be anomalous—it is much further from the line than any other point; indeed, it is higher than the points for the next two larger masses, so it ought to be checked. In fact, the point is incorrect, as a result of a simple arithmetic slip (Figure 9). Because the unextended length of the wire was recorded, this is an easy mistake to spot and to rectify; but if the extensions had been written down directly (without the 'zero reading') this kind of error would have been impossible to trace.

20.0	2.2	1.0
22.5	2.4	1.5 1.2
25.0	2.5	1.3

FIGURE 9 Correcting the record.

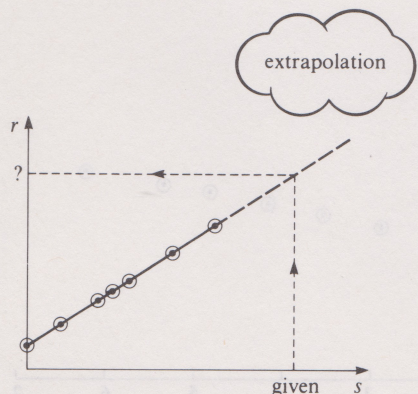
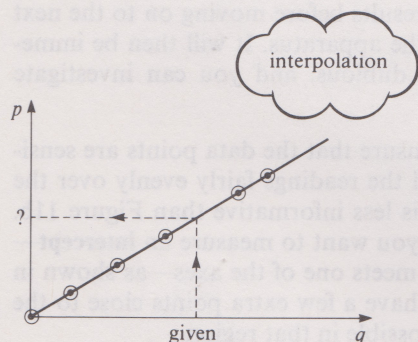


FIGURE 10 Interpolation and extrapolation of straight-line graphs.

- ☐ What extension would you expect for a mass of 22 kg?
- A mass of 22 kg would be expected to produce an extension of 1.1 mm. This kind of procedure—reading *between* data points on a graph—is called **interpolation**. The processes of extrapolation and interpolation are both shown in Figure 10.
- ☐ What is the *average* increase in length of this wire for an increase in load mass of 5 kg (up to a maximum load of 30 kg)?
- We could take all the data from Figure 7, calculate the increases in

INTERCEPT

length for each additional 5 kg load mass and take the average. Merely describing such a procedure is probably enough to convince you that it would be a rather tedious task! Fortunately, the graph saves us from most of this labour by doing the averaging for us. According to the best-fit line in Figure 8, the total extension corresponding to an increase in mass from zero to 30 kg (i.e. virtually the full range of the straight-line portion) is 1.54×10^{-3} m or 1.54 mm. So the average extension is (1.54/30) mm per kg of load mass, and the average extension for an increase of 5 kg in the load mass is $(1.54 \times 5/30)$ mm $\approx$ 0.26 mm. We will return to the averaging function of graphs in the AV sequence (Section 5).

All this information *could* have been extracted from the tabulated data—after all, the graph was plotted solely from the information given in Figure 7—but this would have been laborious. Graphs make relationships more obvious, which is why scientists are so keen on drawing them. The form of the relationship between measured quantities, the typical uncertainties in the measurements (as indicated by the scatter of the points about the best-fit line), and the presence of anomalous measurements are more readily apparent. Furthermore, graphs allow straightforward averaging of experimental measurements, interpolation between measurements and, often, extrapolation beyond the experimental range.

ITQ 2 As a brief exercise in ‘creative scientific thought’, write down a short list of the sorts of experiments you consider it would be useful to try as the next step in your research into the behaviour of copper wires under load. (This question allows plenty of scope for curiosity and imagination, but don’t spend *too* long on it!)

The example of Figures 8 and 9 illustrates how an apparently anomalous point on a graph can be traced back to a simple arithmetic mistake. However, the point might have been a long way from the best-fit line for some quite different reason, such as a misreading of a length. To spot these sorts of mistakes, it is good practice, whenever possible, to plot a rough graph as the experiment proceeds. If the experiment doesn’t lend itself to this, then you should certainly plot the results before moving on to the next stage of the experiment or dismantling the apparatus. It will then be immediately obvious if any data points are dubious, and you can investigate apparent anomalies at once.

Plotting as you go along also helps to ensure that the data points are sensibly spaced. It is generally best to spread the readings fairly evenly over the full range available to you; Figure 11a is less informative than Figure 11b. One exception to this rule occurs when you want to measure an **intercept**—the point at which the line of the graph meets one of the axes—as shown in Figure 12. In this case it is desirable to have a few extra points close to the axis, to make the graph as accurate as possible in that region.

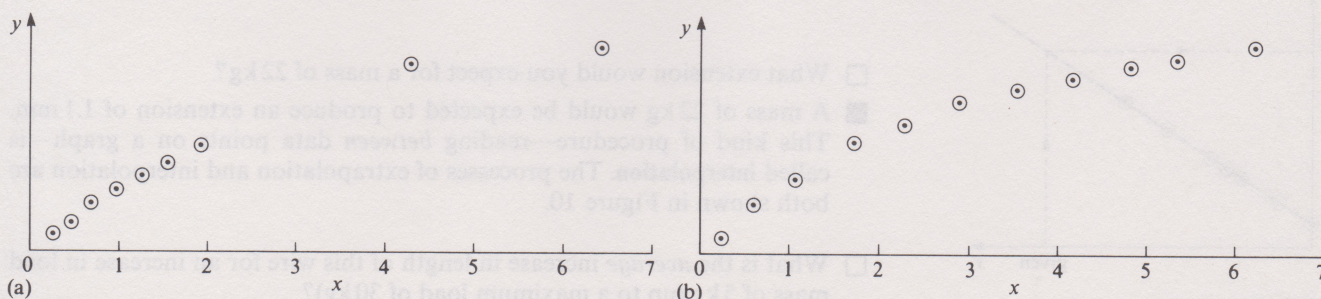


FIGURE 11 Choosing the spread of data points: (b) represents better experimental practice than (a).

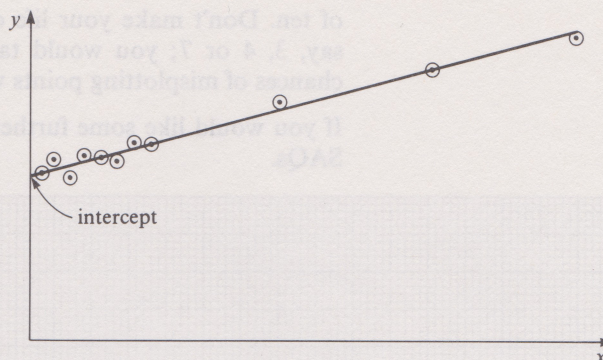


FIGURE 12 Measuring an intercept: extra points near the axis are helpful.

Allied to a sensible spread of data points is the choice of a suitable scale for your graph: the divisions on the axes should be chosen so that the points are well spread across the paper, not cramped up in one small area. In some cases this may mean excluding zero from one or both of the axes. For example, if lengths measured in an experiment varied between 5.4 m and 7.7 m, it might be better to run the scale from 5 m to 8 m as in Figure 13b, rather than from zero to 8 m as in Figure 13a. Be careful, though! Graphs on which the scales are very compressed or very extended can be misleading. Figure 14 illustrates this process carried to extremes.

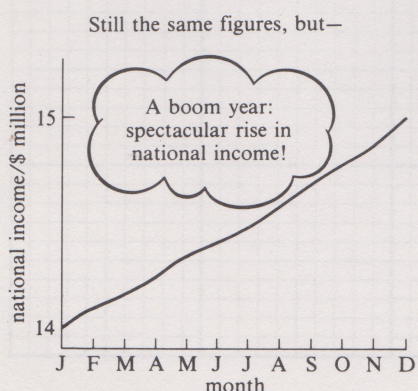
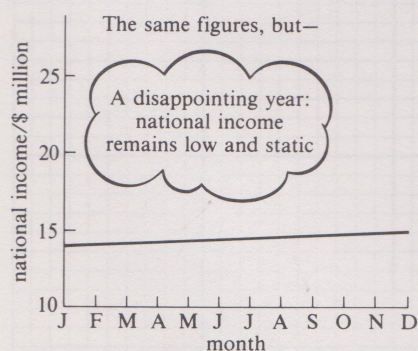
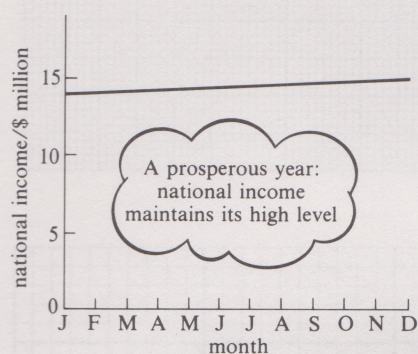


FIGURE 14 Manipulating the statistics.

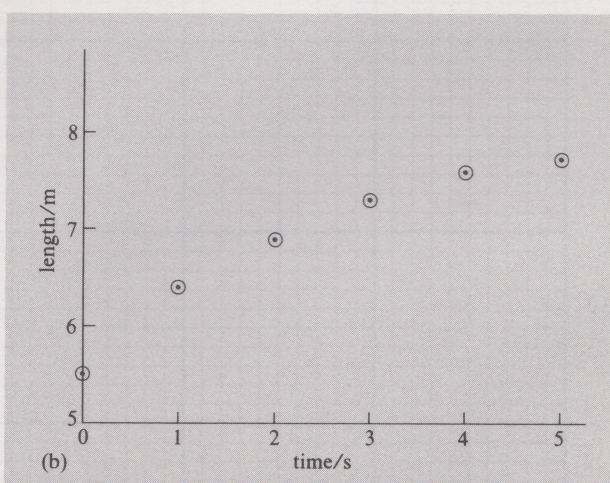
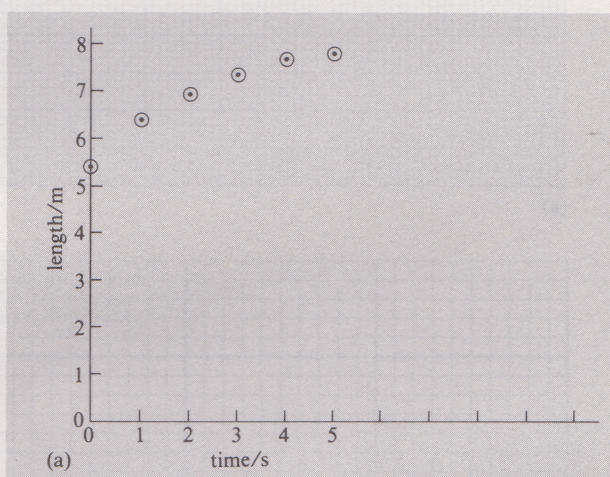
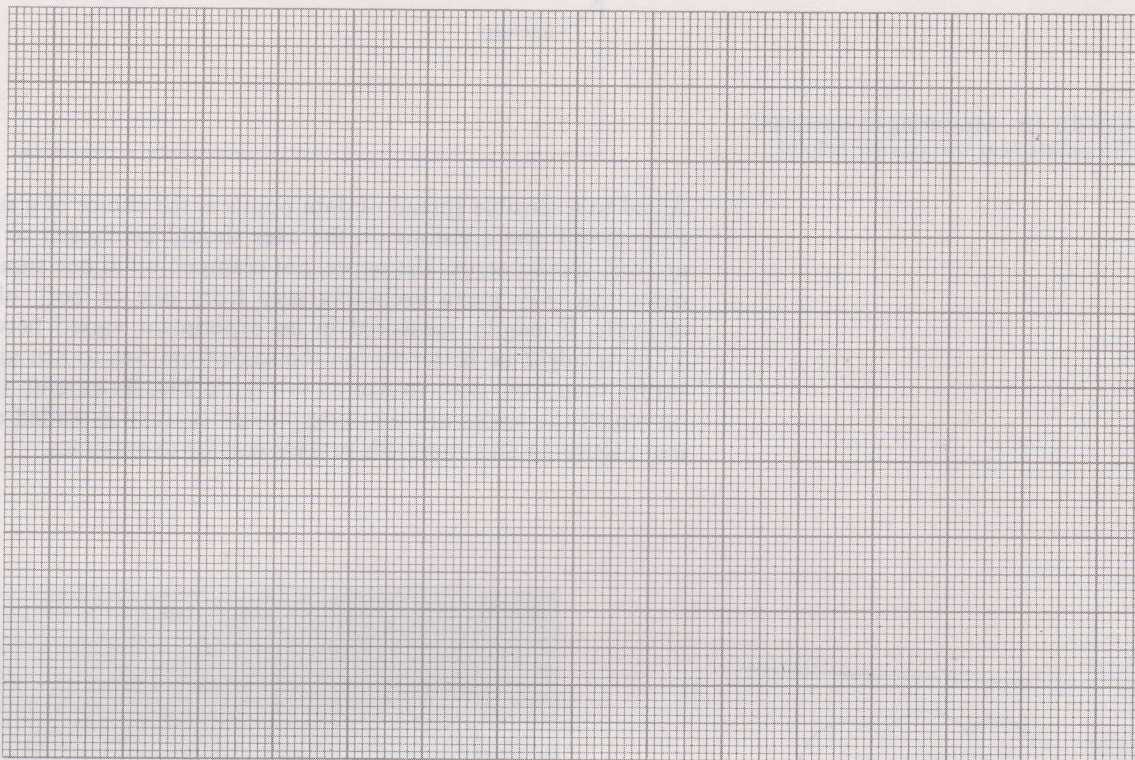


FIGURE 13 Choosing a scale: (a) is rather a poor choice, (b) is more sensible.

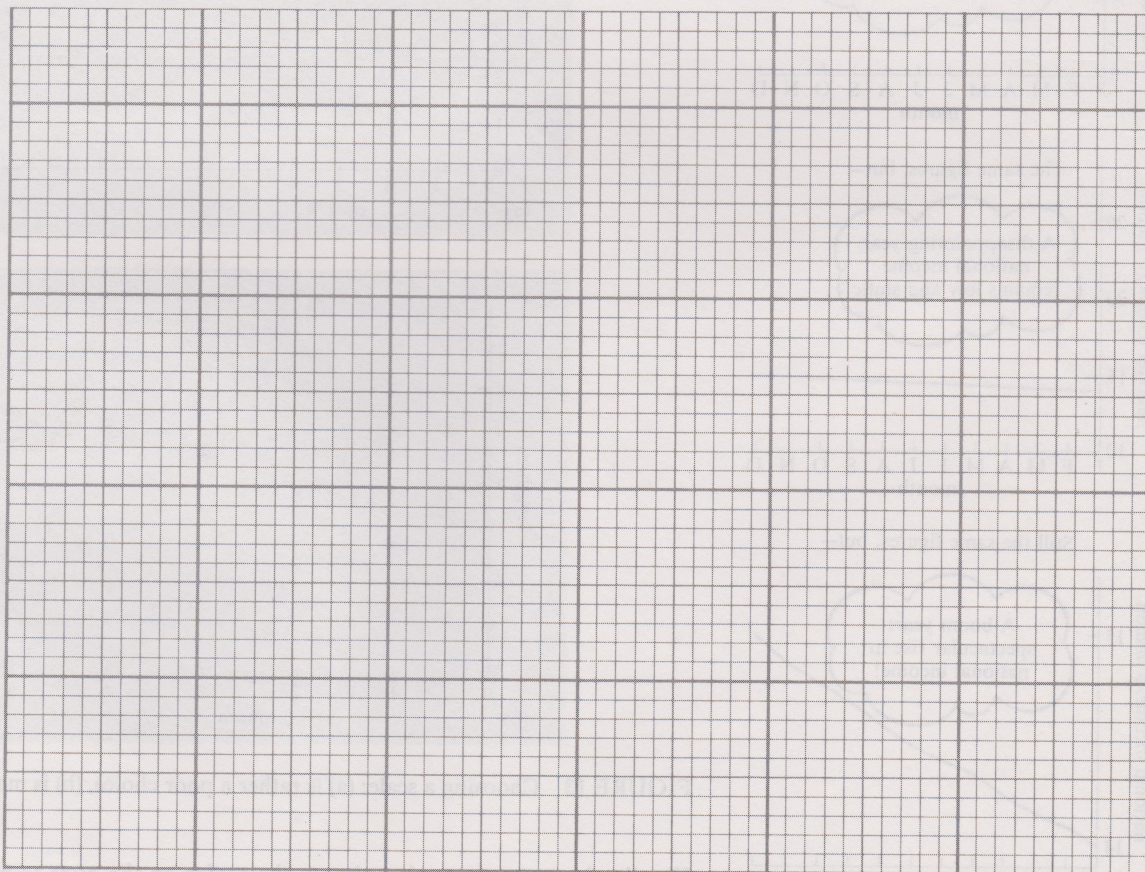
Finally, for your own sake, choose the scales on the axes to make plotting as simple as possible. Generally this means letting ten small divisions on the graph paper equal 1, 2 or 5, or one of these numbers multiplied by a power

of ten. Don't make your life difficult by making ten small divisions equal, say, 3, 4 or 7; you would take much longer to plot the graph, and the chances of misplotting points would be much higher.

If you would like some further practice at graph-plotting, try the following SAQs.



(a)



(b)

FIGURE 15 For use with SAQ 1.

TABLE 1 Heart rate (R) measured as a function of the time (t) elapsed since exercise finished

t /minutes	R /beats per minute
0	180
1	166
2	152
3	138
4	124
5	107
6	98
7	90
8	85
9	80
10	78

TABLE 2 Information taken from the Highway Code, showing the braking distance for a car as a function of its speed

Speed/m.p.h.	Braking distance/ft
20	20
30	45
40	80
50	125
60	180
70	245

TABLE 3 Speed of car

Time/s	Speed/ km h^{-1}
0	64
10	69
20	73
30	76
40	78
50	79

SAQ 1 One of the main skills required in drawing graphs is the ability to choose the appropriate scales for the axes, taking into account both the range of values to be plotted and the type and size of graph paper available.

(a) Present the data in Table 1 as a graph, using the graph paper in Figure 15a and plotting time along the horizontal axis.

(b) Present the data in Table 2 as a graph, using the graph paper in Figure 15b and assuming speed to be the independent variable. Use your graph to find the braking distance for a car travelling at 45 m.p.h. (Notice that for once the units are not SI, indeed not even metric; in real-life situations, and certainly where the traffic laws are concerned, we do have to be able to handle other systems of units.)

SAQ 2 Having noted the speed of a car at various times, as given in Table 3, an inexperienced graph-plotter presents the measurements as shown in Figure 16a. This graph can be criticized on at least six counts. List all the shortcomings you can see on Figure 16a, and correct them by replotting the results on the graph paper of Figure 16b. Hence estimate the speed of the car after 25 seconds.

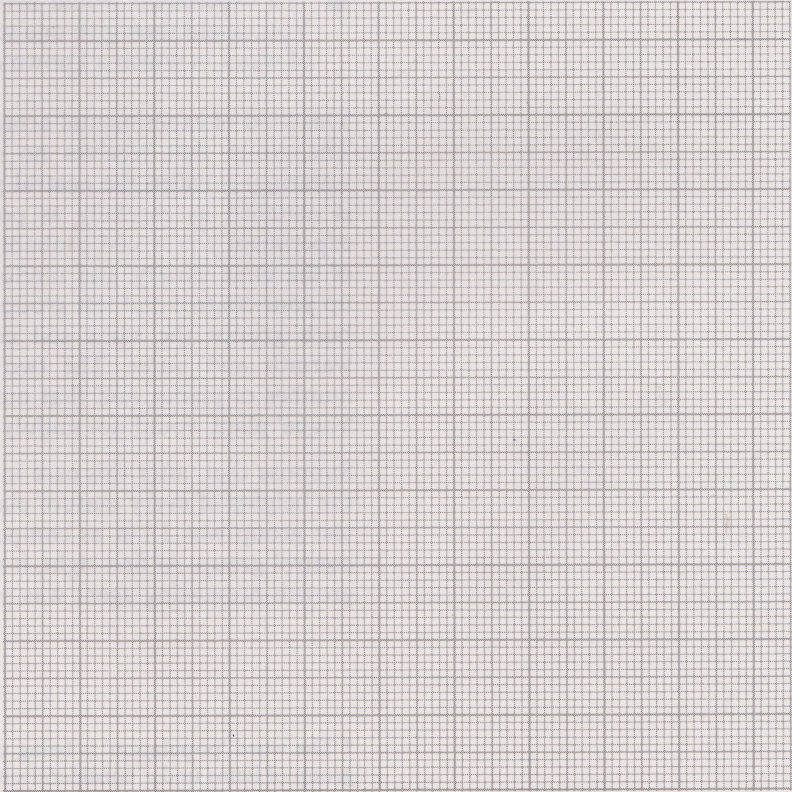
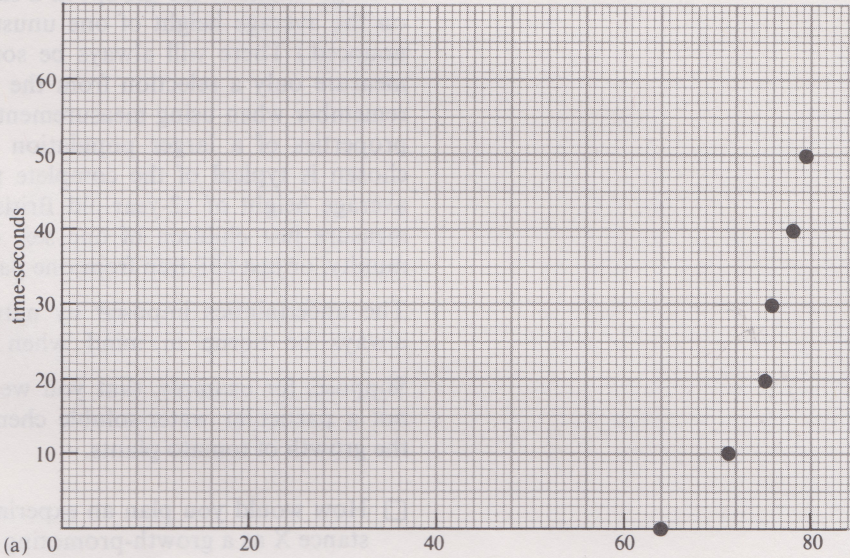


FIGURE 16 For use with SAQ 2.

(b)

CONTROL GROUP

3.2 IDENTIFYING EXPERIMENTAL UNCERTAINTIES

Uncertainties in experimental data arise from a wide variety of causes. For example, measurements on populations of living organisms are, as you saw in Section 2.2, particularly prone to two types of uncertainty—one due to natural variation between individuals in the sample, and the other due to the choice of an unrepresentative sample.

If we measured the heights of ten 12-year-old children, we would find a range of values from which we could calculate the average height of these children. If we were to repeat this with another sample of ten 12-year-olds, the average height would probably be different. So, although measurements on a sample of ten children give us an *estimate* of the average height of 12-year-olds, it isn't a precise value—there must be some uncertainty associated with the fact that we have only taken a limited sample. As you might expect, the uncertainty in the estimate of the average height of the population* becomes smaller as the size of the sample increases. (One very tall individual in a sample of ten children has a significant effect on the average height of the group; in a sample of a thousand, however, the effect on the average height of one unusually tall individual would be less pronounced.) There will always be some statistical uncertainty whenever we measure only a selection from the complete population. Another point to remember when using measurements made on a small group to deduce the properties of a larger population is that it is essential that the sample chosen is typical of the complete population. If we wanted to know the average height of 12-year-old British children, then it would be wrong to measure just children of one sex, or just children from one ethnic community, or just children from one particular locality.

The uncertainties imposed by natural variation within the sample must always be borne in mind when analysing data on living organisms.

Suppose, for example, that you were interested in finding out whether or not a particular water-soluble chemical—let's call it substance X—affects the growth of tomato plants.

- ☐ How would you plan an experiment to investigate the potential of substance X as a growth-promoting agent?
- ☒ The experiment would involve taking two groups of tomato seedlings. You would then give one set pure water and the other the same amount of water in which a known quantity of substance X had been dissolved.

This kind of experimental arrangement is quite common in biological work. The group that receives no special treatment (in this case just normal watering) is called the **control group**. It is important that the control group is matched as carefully as possible to the test group. In this particular experiment you would probably start by sowing a large number of seeds, all of the same variety, and randomly assigned to one of two batches to make up your control and test groups. You would have to be careful about giving both groups the same soil conditions, and the same amounts of light, warmth and water. The two groups should contain roughly the same number of plants, and that number should be fairly large, so that the odd 'rogue' plant would not unduly bias your results.

Suppose that, after this care in the design of the experiment, you measured the heights of every mature plant, and found that the average height of the test group was 1.1 m as compared with an average height of 1.0 m for the control group. At first sight, it might seem that the case for substance X as a growth-promoting agent was proved. However, you then look more closely

* The word 'population' is being used in its statistical sense here, meaning the total group of items under consideration.

at your data and notice that the heights of individual plants in the control group range from 0.5 m to 1.3 m, while those in the test group vary from 0.6 m to 1.2 m. Obviously, these two ranges overlap considerably. Maybe substance X has no real effect after all, and it is simply by chance that the test-group plants have a slightly greater average height than the control-group plants. How can you be sure that the difference in the averages is really due to the 'X factor'? Fortunately, there are well-tried methods for coping with this sort of problem—the 'statistical significance tests'. As the name implies, these are based on statistical techniques for establishing how meaningful the results are. For the moment, you need know nothing about such tests other than the fact that they exist. Later in the Course, though, you will have an opportunity to see how significance tests are applied in practice when you analyse the data you collect in the biology lab at Summer School.

Meanwhile, we will turn our attention to the sort of experiment in which the quantity we are trying to measure has a unique and definite value (a measurement of the magnitude of the acceleration due to gravity at a particular place, for example), and in which the aim is to determine this value as precisely as possible. What factors affect the uncertainty in the measurement? Broadly speaking, such factors fall into two categories:

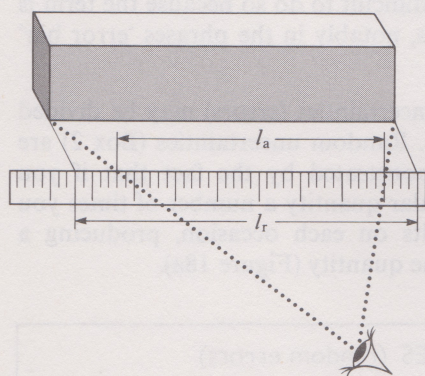


FIGURE 17 How parallax errors arise when the length of an object is measured. The apparent length, l_a , is in this case shorter than the real length, l_r , which would be measured if the ruler were placed in contact with the object.

1 *Human limitations* Often these are simply due to a lack of skill on the part of the experimenter. Parallax errors (Figure 17) can be included in this category, as can the errors introduced by misaligning a measuring instrument or not setting it up as accurately as possible. In general, all these problems tend to be reduced as one becomes more experienced at experimental work. Also, modern instruments are designed to minimize such errors. For instance, meters with digital (i.e. numerical) displays are increasingly replacing needle-and-scale ones, which are prone to parallax errors.

2 *Instrumental limitations* All measuring instruments have their limitations. Some have obvious faults; for example, one can see by eye that the divisions on some cheap wooden rulers are not equally spaced. Other limitations are not so apparent, but even the most expensive and sophisticated instrument will only be capable of measuring to within a given uncertainty limit, and this limit is generally specified by the manufacturer. The manufacturer may also say that the instrument's maximum reliability will only be obtained under certain specified conditions. For example, the divisions on a metal ruler may have been engraved with the ruler at 20 °C; measurements made with this ruler will become less reliable the further the temperature deviates from 20 °C, owing to the expansion or contraction of the metal. In addition to limitations imposed by the design and manufacture of the instrument, there are generally limitations associated with the act of making the readings. These limitations are determined partly by the fineness of any scale divisions, and partly by how well the observer can interpolate between scale divisions. (Note that interpolation is not possible with a digital display.) Another important source of instrumental error arises from unavoidable imperfections in the apparatus, such as friction in the balancing mechanism of a scale.

3.3 ACCURACY AND PRECISION IN EXPERIMENTAL WORK

The inescapable conclusion is that it is never possible to make a measurement without there being some degree of uncertainty associated with the result. The remainder of Section 3 is concerned with making quantitative estimates of such uncertainties and incorporating them into calculations and graphs.

ACCURATE MEASUREMENT

PRECISE MEASUREMENT

CALIBRATION

We have already discussed experimental uncertainties at some length in both Units 2 and 3. At this stage, it is worth reviewing the ground we have covered so far. The starting point is summarized in Box 1.

BOX 1 UNCERTAINTIES

No measurement is ever exact: there is always some *uncertainty* associated with the act of measurement. When reporting the results of practical work, it is important to put a value to the uncertainties, so as to indicate the confidence limits that may be attached to the final results.

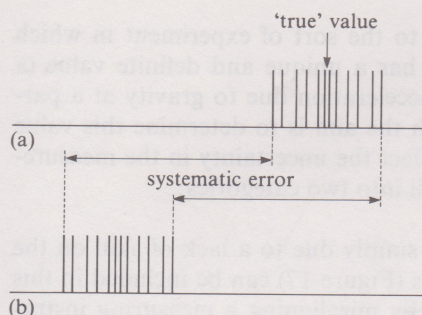


FIGURE 18 (a) Random uncertainties on their own lead to a spread of measurements about the true value. (b) Systematic errors displace all the measurements by the same amount in the same direction.

'Uncertainty' is the best word to use when we know only that the true value of some measured quantity lies within a certain range. We can set upper and lower limits on the quantity, but we cannot be certain of its true value. It is slightly confusing that, when discussing data analysis, many scientists use the words 'uncertainty' and 'error' as synonyms. In this context, the word 'error' does *not* carry any implication of mistake; a phrase such as 'with an error of $\pm 10\%$ ' is synonymous with 'with an uncertainty of $\pm 10\%$ ' (or 'within a confidence limit of $\pm 10\%$ '). It would certainly be clearer to drop the word 'error' altogether in connection with uncertainties in experimental measurements, but it is difficult to do so because the term is in such widespread use among scientists, notably in the phrases 'error bar' and 'systematic error'.

You saw in Unit 3 that experimental uncertainties (errors) may be divided into two types: random and systematic. Random uncertainties (Box 2) are present in every measurement, as demonstrated by the fact that if you repeat the measurement of some particular quantity a number of times you will usually get slightly different results on each occasion, producing a spread of data about the true value of the quantity (Figure 18a).

BOX 2 RANDOM UNCERTAINTIES (random errors)

These arise from random operating errors or from limitations in the scale on which the reading is based.

You have seen in both Units 2 and 3 that we can, to some extent, 'cancel out' the effects of random uncertainties by taking the average of a number of readings, or drawing the best line through the data points plotted on a graph. Furthermore, the spread of the readings (or the scatter of the points about the line of best fit on the graph) gives some indication of the size of the random uncertainties involved.

Systematic errors (Box 3) are often superimposed on the inevitable random uncertainty associated with every measurement.

BOX 3 SYSTEMATIC ERRORS

These are said to have arisen when *all* the measurements have been shifted in the same direction away from the true value.

Systematic errors are not usually shown up by repetition of the measurements, and averaging does nothing to reduce their effects. Typically, they arise from problems such as failure to zero an instrument correctly, 'drift' in readings due to a continual increase in temperature, or progressive contamination of a sample. The combined effects of random and systematic errors are illustrated in Figure 18b.

In scientific terminology, there is an important distinction between accuracy and precision. An **accurate measurement** is one in which the systematic errors are small. A **precise measurement** is one in which the random uncertainties are small. Ideally, the results of your experiments should be both accurate and precise!

3.4 ESTIMATING EXPERIMENTAL UNCERTAINTIES

Often one of the most difficult parts of an experiment is the estimation of the uncertainties in the measurements. Because no two experiments are exactly alike, it is impossible to give hard-and-fast rules about how to tackle this aspect of practical work. Instead, we shall consider a few examples that illustrate methods of quantifying uncertainties. You will then have to make up your own mind about how to estimate the experimental uncertainties in each piece of practical work you carry out. Don't worry though—as you will see, these estimates are largely just a matter of common sense.

In some ways, systematic errors are more difficult to deal with than random ones, because they often arise from some problem with the measuring instrument that is hard to detect. For example, a metre rule may in fact be 1.005 m long (i.e. 5 mm overlength), even though all its graduated divisions are equal. All measurements made with this particular rule would be consistently (i.e. systematically) 0.5% too short (since 0.005 in 1.000 is equivalent to 0.5 in 100, i.e. 0.5%). Systematic errors, like these, in measuring devices can often only be discovered and estimated by comparison with a more accurate and reliable instrument (a process known as **calibration**), and this should be done whenever possible. The difficulty with systematic errors lies in detecting them in the first place; once you are aware of such an error, it can often be eliminated by making some modification to the apparatus or by re-calibrating it. Failing that, at least a correction can be made to the final result.

Procedures for estimating random uncertainties are rather different, and you have already had some practice in this during your work on Units 2 and 3. As you have seen, the first step consists of seeing whether your measurements can be exactly repeated. Usually, you should not be satisfied with a single measurement: repeat every measurement several times if possible, resetting the measuring instrument each time if you can. Thus, if you were measuring the length of a rod, you would remove the ruler, reposition it and read off the length of the rod, then repeat this procedure several times.

If you measure a particular quantity several times but get slightly different answers each time, then it is quite easy to ascribe a value to the uncertainty. Suppose, for example, that you have been measuring the flow rate of water through a pipe by collecting the water in a measuring jug and recording the volume collected over a four-minute interval. You have made five such measurements and your results are:

$$436.5 \text{ cm}^3, \quad 437.5 \text{ cm}^3, \quad 435.9 \text{ cm}^3, \quad 436.2 \text{ cm}^3, \quad 436.9 \text{ cm}^3$$

Assuming that all these measurements are equally reliable (i.e. they were all taken with the same care and with the same measuring instruments), then the best estimate of the water volume is the average value of the five readings. So the best estimate of the volume is:

$$\frac{(436.5 + 437.5 + 435.9 + 436.2 + 436.9) \text{ cm}^3}{5} = 436.6 \text{ cm}^3$$

Now, what error should we associate with this average value? The measurements are spread out between 435.9 cm^3 and 437.5 cm^3 , i.e. from $(436.6 - 0.7) \text{ cm}^3$ to $(436.6 + 0.9) \text{ cm}^3$. If we average these deviations of 0.7 and 0.9 cm^3 , then we can reasonably say that the spread is about $\pm 0.8 \text{ cm}^3$.

Admittedly, the true volume is unlikely to lie right at the ends of this spread, so quoting an uncertainty of $\pm 0.8 \text{ cm}^3$ may be unduly pessimistic. However, this is not necessarily a bad thing. All you are really doing is putting a confidence limit on your result, and you can be reasonably confident on the basis of these data that the true value lies within $\pm 0.8 \text{ cm}^3$ of the average of your readings. There are more sophisticated methods, based on statistical theory, for working out uncertainties from repeated measurements of the same quantity, and Section 4 looks at one of these statistical techniques. However, in nearly all the cases that arise in this Course, the straightforward procedure outlined above is perfectly adequate.

SIGNIFICANT FIGURES

You will appreciate that if one reading is very different from all the others, then a determination of the likely error from the spread will probably be misleadingly pessimistic. On the whole, such odd readings should be ignored when calculating the average and the uncertainty, and ideally a few more measurements should be taken.

In our example of water flowing through a pipe, the variability in the measured volumes could have arisen from the combination of a number of factors. These include actual variations in the flow rate, errors in timing four-minute intervals, inserting and removing the jug too early or too late, and inaccuracies in measuring the volume collected. The point of repeating the measurements is that you then get an overall measure of the random uncertainties involved, so it isn't necessary to assess their individual contributions.

You will find when repeating some experimental measurements that you record exactly the same result each time. Does this then mean that there is no uncertainty associated with the measurement? In most cases, it certainly does not! Suppose you had been measuring some particular length and, on repeating the measurement three times you had obtained:

73 mm 73 mm 73 mm

that is, three identical readings. You have no spread in your results to help you assess the uncertainty in your measurement *but that does not entitle you to conclude that the uncertainty is negligible.*

If someone else were to do the same experiment, perhaps with different equipment, they would still want to know what confidence limit you placed on your result, so as to see whether their answer was consistent with yours. For example, suppose their value for the length was (72.4 ± 0.4) mm and yours was (73.0 ± 0.3) mm; these two sets of results are consistent, since their ranges overlap. Separate experiments giving results of (72.4 ± 0.2) mm and (73.0 ± 0.1) mm, on the other hand, should not be said to be in agreement.

It is, therefore, important always to state a confidence limit with every experimental result. If there is no spread in your readings of a particular quantity, or if, for some reason, you can make only a single reading, you must look at your equipment to see what limitations it imposes on your precision. (Of course, you should do this for *every* experiment, but in these situations it is the *only* method of assessing the reliability of your result.)

The precision you can claim for a measurement will always depend on the scale divisions on the instrument you are using. Figure 19 shows two rulers being used to measure the length of a rod. With the top ruler, a measurement to the nearest 0.1 mm could probably be obtained by interpolation; 72.8 mm would be the best result one could give. The bottom scale, however, can be read to a greater degree of precision; a value of 72.85 mm could reasonably be recorded using this ruler.

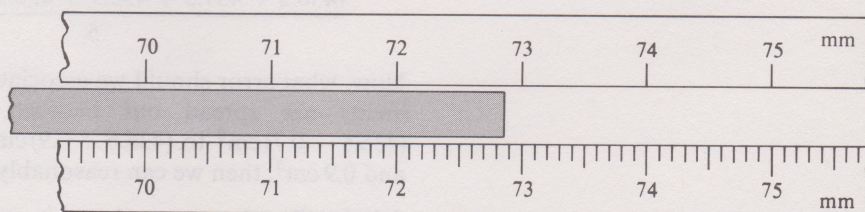


FIGURE 19 The top scale can be read only to within 0.1 mm at best, but the bottom scale can be read to within 0.02 mm.

3.5 QUOTING EXPERIMENTAL RESULTS: UNCERTAINTIES AND SIGNIFICANT FIGURES

The number of figures that you give when quoting data is not arbitrary: it indicates the precision with which the measurement has been made. Even if no uncertainty is explicitly given, it is implicit in the number of digits quoted in the result. Thus, if you use the top ruler in Figure 19, and write down your measurement as 72.8 mm, you are implicitly stating that the length of the rod is closer to 72.8 mm than it is to 72.7 mm or to 72.9 mm—in other words, that the length lies between 72.75 mm and 72.85 mm. You are therefore claiming an uncertainty of ± 0.05 mm, and should quote your result as (72.8 ± 0.05) mm. Using the bottom ruler in Figure 19, you could probably interpolate the scale with an uncertainty of ± 0.01 mm; if so, you would be justified in recording your answer as (72.85 ± 0.01) mm. Be careful, though, in interpolating between scale divisions, and do not claim a level of precision that you cannot really achieve!

These are examples of an extremely important general rule: *never quote a result to more digits than you can justify in terms of the limitations of your equipment.* You have already met this idea of ‘significant figures’ in both Units 2 and 3, but the importance of this concept cannot be over-emphasized (Box 4).

BOX 4 SIGNIFICANT FIGURES

The accurately known digits in the value of a physical quantity, plus one uncertain digit, are called **significant figures**. Experimental results should always be quoted to a number of significant figures consistent with the known uncertainties in the measurements. If two or more quantities are combined (e.g. by dividing one by the other), then the result is known only to the same number of significant figures as the *least* precisely known quantity.

So, if your calculator displays the result of a calculation as 2.0375, you could write this as

2	(to one significant figure)	2.04	(to three significant figures)
2.0	(to two significant figures)	2.038	(to four significant figures)

Note the convention: if the last significant figure would have been followed by a number from 0 to 4, it is unchanged, but if it would have been followed by a number from 5 to 9, it is rounded up.

ITQ 3 Use your calculator to work out, in km s^{-1} , and to an appropriate number of significant figures, the speeds corresponding to the following data:

- A car covers a distance of 1 km in 50 s.
- A lorry covers a distance of 1.20 km in 66 s.
- Light covers a distance of 3000 km in 0.01 s.

(Make sure you read the answer and comment to this ITQ before proceeding.)

ITQ 3 demonstrates how useful scientific (or powers-of-ten) notation can be in this context: it avoids any confusion about whether to count zeros as significant figures.

The correct use of significant figures also reinforces the idea that when several measurements are combined it is the *least* precisely known quantity that determines the overall uncertainty of the final result. When doing experiments, it is therefore very important not to waste time reducing small errors or uncertainties when much larger errors or uncertainties are also present. As a general rule, *find out as early as possible in an experiment what the dominant uncertainties are, and then concentrate your time and effort on reducing them.*

ERROR BAR

3.6 GRAPHICAL REPRESENTATION OF UNCERTAINTIES

Once you've taken the trouble to estimate the uncertainties associated with your data, it is a simple matter to transfer this information onto a graph.

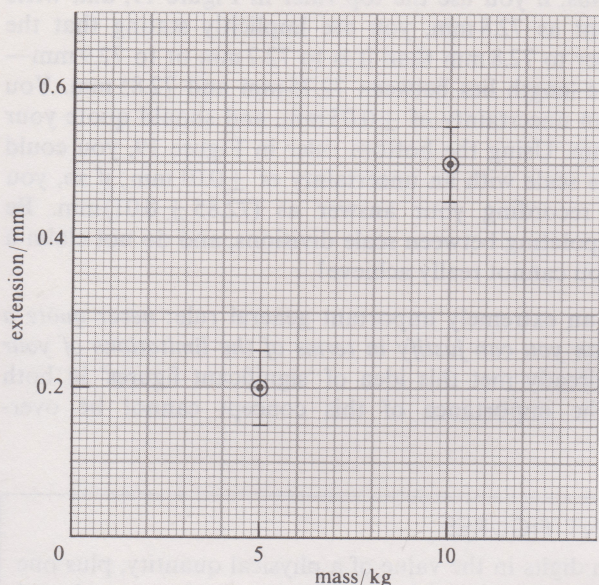


FIGURE 20 Vertical error bars indicate that the extension is known to ± 0.05 mm. The absence of horizontal error bars implies that the mass is known much more precisely.

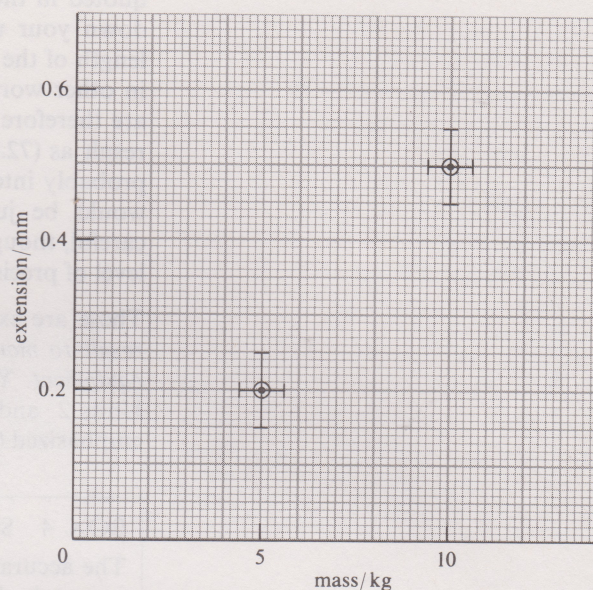


FIGURE 21 Vertical and horizontal error bars. The horizontal bars indicate an uncertainty of ± 0.6 kg.

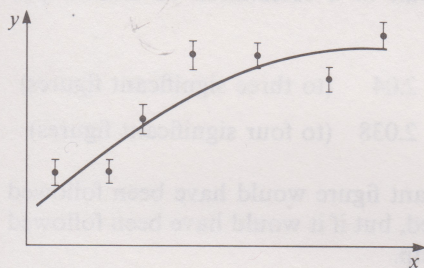


FIGURE 22 Have the error bars been underestimated?

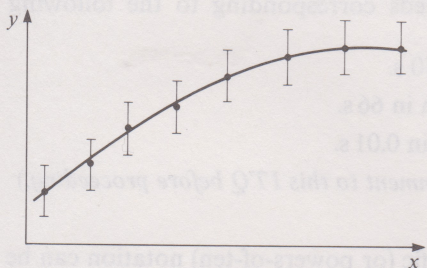


FIGURE 23 Have the error bars been overestimated?

For example, if the extensions measured in the wire-stretching experiment (Figure 7) are accurate to ± 0.05 mm, then the first two measurements would be represented graphically by points and vertical bars, as shown in Figure 20. These **error bars** (i.e. uncertainty limits) extend 0.05 mm above and below the plotted points, which are indicated by dots with circles round them. The fact that there are no horizontal error bars shows that it has been assumed that the masses are known much more accurately than the extensions. In fact, they must have an uncertainty of less than ± 0.2 kg because this is the smallest uncertainty that could be shown on the graph, given the size of the circles around the data. Suppose, however, that the masses were known only to within ± 0.6 kg. This uncertainty would be represented by a horizontal error bar extending 0.6 kg either side of the plotted point, as shown in Figure 21. Generally, both horizontal and vertical error bars should be shown, and either of them omitted only if the associated uncertainty is too small to show up on the graph.

Error bars on graphs serve a number of useful purposes. If we predicted that a particular graph would be a smooth curve, then we would expect the points to be scattered around that curve by amounts ranging up to the size of the error bars. If the points deviated from a smooth curve by much more than the error bars, as shown in Figure 22, then we should suspect either that we had underestimated the errors, or that our assumption that a smooth curve would describe the results was not valid. On the other hand, if all the points deviated from the expected curve by much less than the error bars, as in Figure 23, then we might suspect an overestimate of the errors.

Finally, just as a graph can do the averaging for you, so it can take care of the uncertainty calculations. You used a graph in this way in Unit 3 when, for the stroboscopic experiment described in the AV sequence, you plotted the vertical position of the ball against the square of the time it had been falling. The graph had roughly the form shown in Figure 24a. Usually, the best-fit straight line through the points can be drawn surprisingly accurately by eye, providing that you remember the simple rule of having

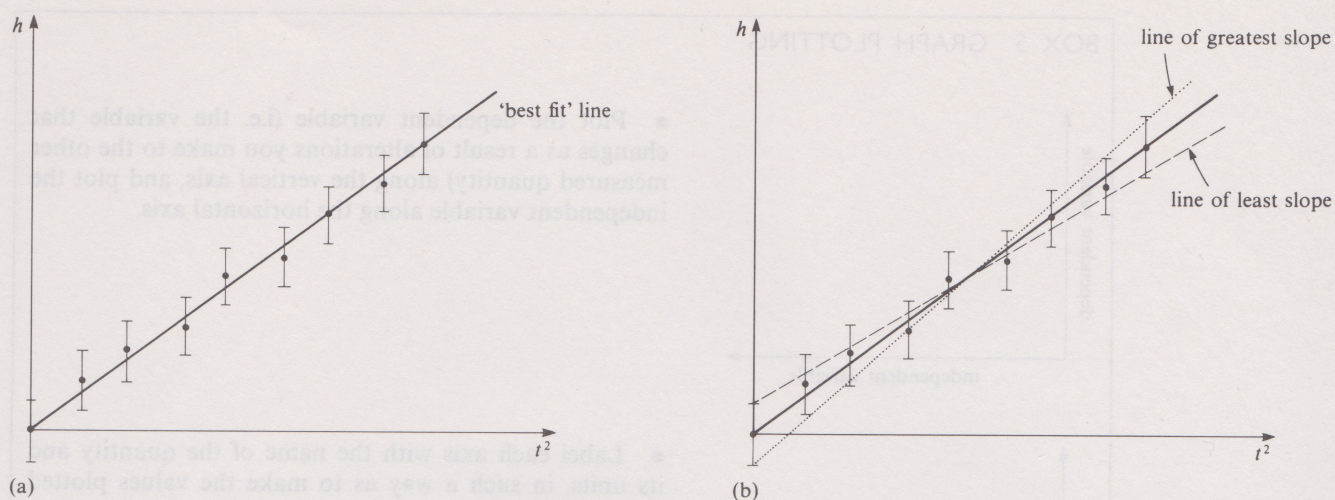


FIGURE 24 Making use of error bars.

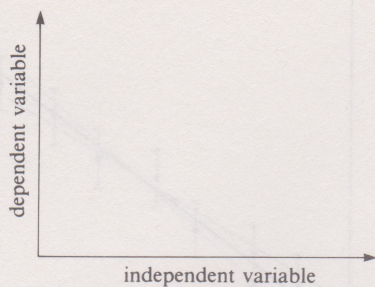
approximately the same number of points scattered above the line as below it. In Unit 3, we related the average or best estimate value of g_E (the magnitude of the acceleration due to gravity at the surface of the Earth) to the slope (*gradient*) of this line. To determine the uncertainty in this value of g_E , all we had to do was to draw the lines of least and greatest slope that still passed through all the error bars (Figure 24b). From the gradients of these lines we could obtain the upper and lower limits on our measurement of g_E .

SUMMARY OF SECTION 3

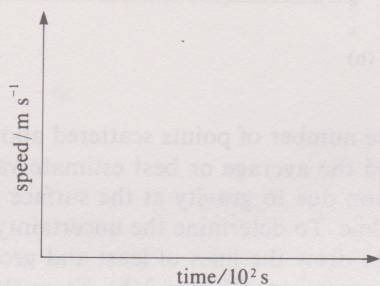
The following summary should serve as a checklist of experimental good practice. With time, all these techniques and precautions will become second nature to you. Meanwhile, use the list as a handy reminder.

- 1 Enter all your data and comments into your Notebook as you carry out the experiment. Record *everything*.
- 2 Plan your experiment in advance; whenever it is feasible, carry out a rough-and-ready rehearsal. When you are actually doing the experiment, however, be flexible: don't miss opportunities to improve the apparatus, or to refine your data. If at all possible, plot graphs as you do the experiment—a graph will help to alert you to potential problems in time to do something about them. Box 5 (overleaf) gives a checklist of points to remember when plotting graphs.
- 3 Try to identify the dominant uncertainty in your data at an early stage, and concentrate your time and effort on reducing it. Uncertainties smaller than about a third of the dominant one will have a negligible effect on the accuracy and precision of your final result.
- 4 When doing calculations, use your common sense. Box 6 gives a checklist of points to look out for.

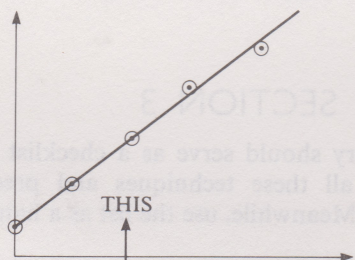
BOX 5 GRAPH PLOTTING



- Plot the dependent variable (i.e. the variable that changes as a result of alterations you make to the other measured quantity) along the vertical axis, and plot the independent variable along the horizontal axis.



- Label each axis with the name of the quantity and its units, in such a way as to make the values plotted dimensionless numbers.

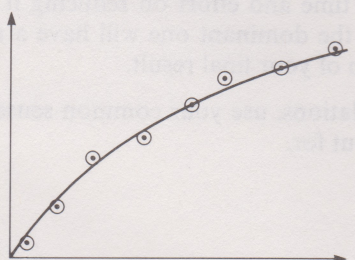
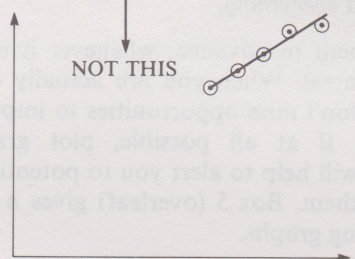


- Choose scales such that the points are sensibly spread over the graph paper.

- Use simple scales that make the points easy to plot.
- Plot points as small dots within circles, or as crosses.

- Include error bars (both horizontal and vertical if there are uncertainties in both quantities).

- Think carefully about whether your line should go through the origin.



- In general, draw a smooth curve through as many points as possible, not a zig-zag.

- Your line of best fit should have roughly as many points above it as below it.

STANDARD DEVIATION

BOX 6 CALCULATIONS

$$(1361 + 75) \times \frac{101}{7} = 20719$$

Check

$$\left(\frac{75 + 1361}{7} \right) \times 101 = 20719$$

OK

$$(1361 + 75) \times \frac{101}{7} = 20719$$

Roughly

$$= 2.0719 \times 10^4$$

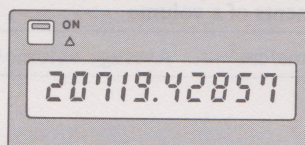
$$(1500) \times \frac{100}{7} \sim 200 \times 100$$

$$\sim 2 \times 10^4$$

OK

Right 'order of magnitude'
(power of ten)

My calculator display shows this: ~

How many of these figures are
significant?

- Check calculations by carrying out the operations in reverse order.

- Check that the order of magnitude of your result is sensible (by using rounded-off numbers).

- Ask yourself whether the result of a calculation or measurement seems reasonable.

- Drop meaningless digits arising from calculations. The final result of a multiplication or division should have no more significant figures than the factor with the fewest.

- Use scientific notation to avoid any confusion about zeros as significant figures. For example, 1.32×10^3 unambiguously has three significant figures, but 1320 would probably be read as having four significant figures.

4 THE STATISTICAL APPROACH: STANDARD DEVIATION

An objective method of reporting the spread in a set of measurements, which has significance for statisticians, is to quote a quantity called the **standard deviation**, s . If you give the spread of your results in terms of the standard deviation, other people will know exactly what you mean.

You do not need to remember the formula for calculating the standard deviation of a set of results. Nevertheless, it is worth seeing the procedure at an early stage in your study of science, because when scientists quote experimental uncertainties, they most often do so in terms of the standard deviation or of a quantity derived from it. An appreciation of the method used to work out standard deviations also brings out a very important point about repeated measurements, as we shall see later in this Section.

The procedure for calculating s can be given in terms of an algebraic formula. To calculate s , you need to know the total number of readings, n , and the individual readings, $x_1, x_2, x_3, x_4, x_5 \dots x_n$. From these you need to calculate the average of the readings, x_{av} , and then the differences between each of the readings and the average, $d_1, d_2, d_3 \dots d_n$. That is

$$d_1 = x_1 - x_{av}$$

$$d_2 = x_2 - x_{av}$$

and so on up to

$$d_n = x_n - x_{av}$$

The standard deviation is then given by

$$s = \sqrt{\frac{d_1^2 + d_2^2 + \dots + d_n^2}{n}}$$

In words, we would say that the standard deviation is the square root of the sum of the squares of the differences between the readings and the average, divided by the number of readings.

As an example, Table 4 shows how to calculate the standard deviation of the five volume measurements that were discussed in Section 3.4.

TABLE 4 Calculating a standard deviation from five measurements of a volume

Volume/cm ³	Average volume/cm ³	d/cm^3	d^2/cm^6
436.5		-0.1	0.01
437.5		+0.9	0.81
435.9		-0.7	0.49
436.2		-0.4	0.16
436.9		+0.3	0.09
2183.0	2183.0/5 = 436.6		1.56

The first step is to calculate the average (436.6 cm³ in this case). Next, calculate the differences, d , between each reading and this average, and then calculate the square of each of these differences. (Note that d^2 is always a positive quantity.) Now add up all the values of d^2 and divide this sum by the total number of measurements, n (five in this case). This gives s^2 . Finally, take the square root of s^2 to get s , the standard deviation. The last step consists of rounding the result off to an appropriate number of significant figures!

From the Table, the sum of $d^2 = 1.56 \text{ cm}^6$

Thus, using the formula for s ,

$$s = \sqrt{\frac{1.56 \text{ cm}^6}{5}} = \sqrt{0.312 \text{ cm}^6} = 0.6 \text{ cm}^3 \text{ (to 1 significant figure)}$$

We could quote the volume as

$$\text{volume} = (436.6 \pm 0.6) \text{ cm}^3$$

You will have noticed that part of the procedure for calculating a standard deviation involves division by the total number of readings, n . As the number of readings increases by a factor of n , the standard deviation (i.e. the uncertainty) decreases by a factor of $\sqrt{n}$. (If you quadruple the number of readings, you halve the uncertainty.) This is in accord with our intuitive feeling that by repeating our measurements several times we should improve the overall accuracy of our result. Indeed, this highlights the major objection to the method we have used up to now of equating the confidence limit to the *spread* of readings: this approximate method gives too much weight to the highest and lowest readings and so produces an unnecessarily pessimistic estimate of the overall uncertainty.

If you have made many readings of a quantity, and recorded a long list of values, it can be difficult to make a quick assessment of the average value and its precision. A more useful way to display the data is illustrated in

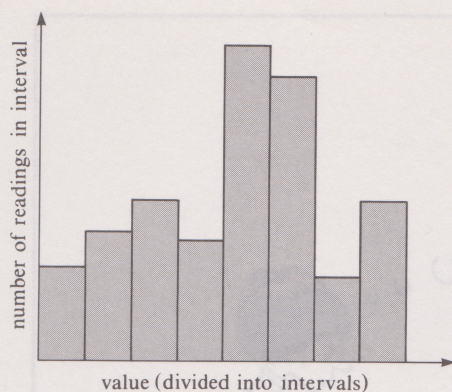


FIGURE 25 Plotting repeated readings of one quantity.

Figure 25. (This kind of graph is called a *histogram*.) Here the range of measured values is divided into equal intervals, which are scaled along the horizontal axis; the number of readings in each interval is plotted along the vertical axis. As the number of readings is increased, so one can take smaller and smaller intervals on the horizontal axis and still have a reasonable number of readings in each interval. Eventually, the step-like character of the graph is no longer noticeable; one simply gets a smooth curve, as shown in Figure 26a.

Provided there is no progressive 'drift' in the readings, any experiment involving a very large number of measurements of one quantity will always produce a bell-shaped curve like the one in Figure 26a. Statistical theory can then be used to relate this curve to the standard deviation, s . It can be shown that about 68% of the readings will lie within $\pm s$ of the true value, 95% within $\pm 2s$, and 99.7% within $\pm 3s$. This is illustrated in Figure 26b.

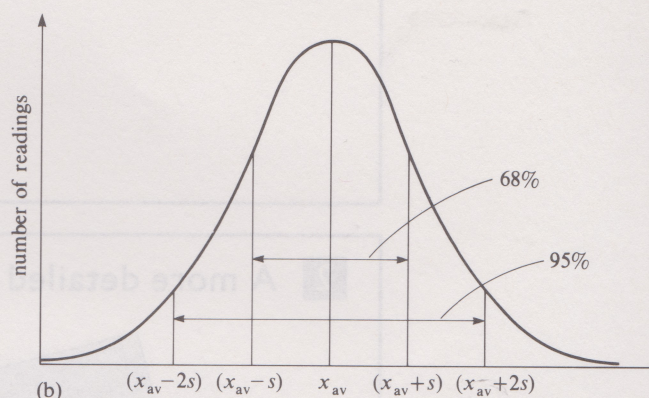
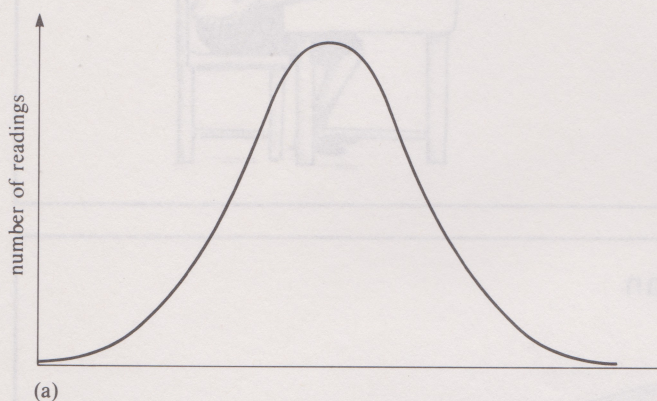


FIGURE 26 Distribution of a large number of readings about an average value.

SUMMARY OF SECTION 4

The standard deviation is related to the spread obtained from repeated readings of a particular quantity. The statistical method used to derive the standard deviation shows that as the number of readings of a quantity increases by a factor of n , the uncertainty in the average value decreases by a factor of $\sqrt{n}$.

5 WRITING REPORTS OF PRACTICAL WORK (AV SEQUENCE)

The time has now come for you to tackle the major part of TMA 01—the report of your determination of the distance from the Earth to the Moon. The aim of this sequence is to help you with this assignment. You'll need to have Unit 2 open at the page describing the experiment. You'll also need plenty of rough paper and your calculator. Most important of all, you must have your Notebook, with your own record of the measurements you made. When you have gathered everything together, switch on the recording, which you'll find on Tape 1 (Side 1, Band 2).

1 The planning stage

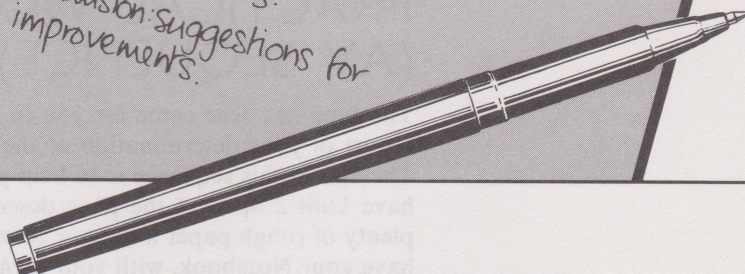
- * Aims
- * Experimental procedure
- * Results
- * Conclusion



2 A more detailed plan

PLAN

- ① Abstract.
- ② Introduction: aims of and background to the experiment.
- ③ Experimental procedure: equipment and method, techniques used to reduce uncertainties, problems encountered, etc.
- ④ Results: raw data/average values
 - uncertainties in measured quantities
 - calculation of final answer
 - confidence limits.
- ⑤ Conclusion: suggestions for improvements.



3 An example of an abstract

The paper announcing the discovery of 'charm'

Discovery of a Narrow Resonance in e^+e^- Annihilation*

J.-E. Augustin,† A. M. Boyarski, M. Breidenbach, F. Bulos, J. T. Dakin, G. J. Feldman,
G. E. Fischer, D. Fryberger, G. Hanson, B. Jean-Marie,† R. R. Larsen, V. Lüth,
H. L. Lynch, D. Lyon, C. C. Morehouse, J. M. Paterson, M. L. Perl,
B. Richter, P. Rapidis, R. F. Schwitters, W. M. Tanenbaum,
and F. Vannucci†

Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305

and

G. S. Abrams, D. Briggs, W. Chinowsky, C. E. Friedberg, G. Goldhaber, R. J. Hollebeek,
J. A. Kadyk, B. Lulu, F. Pierre,§ G. H. Trilling, J. S. Whitaker,
J. Wiss, and J. E. Zipse

*Lawrence Berkeley Laboratory and Department of Physics,
University of California, Berkeley, California 94720*

(Received 13 November 1974)

The
abstract

We have observed a very sharp peak in the cross section for $e^+e^- \rightarrow \text{hadrons}$, e^+e^- , and possibly $\mu^+\mu^-$ at a center-of-mass energy of 3.105 ± 0.003 GeV. The upper limit to the full width at half-maximum is 1.3 MeV.

4 Beginning the report

Introduction

○ Aim

.....

.....

.....

○ Brief outline of the eclipse method
– equation :

.....

– explanation of how quantities are to
be determined :

.....

.....

.....

.....

.....

One
sentence
per
quantity

Ringed terms
to be
defined

Diagram?

5 The next section

Experimental procedure

- ☐ Setting up the apparatus

.....

.....

- ☐ Procedure for determining the various quantities (including any special innovations)

Quantity	Method	Innovations
.....		
.....		
.....		
.....		

- ☐ Were any measurements repeated?

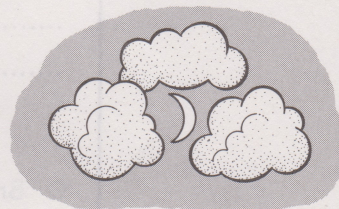
.....

.....

- ☐ Problems (e.g. Moon only intermittently visible, measurements not made on full Moon, etc.)

.....

.....



- ☐ Limitations of the equipment; precautions taken to ensure highest possible accuracy/precision

.....

.....

.....

.....

Reminder:
350–400 words
for this
section

6 The data

(i) From the lunar eclipse photographs:

(ii) From the eclipsing experiment:

7 Working out the results (I)*Calculation of the Moon's radius*

Step 1 Work out $\frac{D_{E(sh)}}{D_M}$ and associated uncertainty:

.....

Step 2 Apply correction:

$$\frac{D_{E(sh)}}{D_M} + 1 = \frac{D_E}{D_M} = \frac{R_E}{R_M}$$

=

Now what is the uncertainty?

.....

Step 3 Given $R_E = 6.38 \times 10^6 \text{ m}$,

Calculate R_M :

.....

What is the uncertainty in R_M as a percentage?

.....

8 Working out the results (II)

Calculation of the distance to the Moon

Step 1 Use $L_M = \frac{D_M l}{d}$

to calculate your 'best value' for L_M :

.....

Step 2 Note the confidence limits on your values of D_M , l and d :

.....

Step 3 Use the simple 'range method' to calculate the maximum and minimum values of L_M possible with these confidence limits:

$$(L_M)_{\max} = \dots\dots\dots$$

$$(L_M)_{\min} = \dots\dots\dots$$

Step 4 Express the results of steps 1 and 3 in scientific notation:

$$L_M = (\dots\dots\dots \pm \dots\dots\dots) \times 10^{\boxed{}} \text{ m}$$

How many
significant
figures?

9 Discussion and conclusion

- Which aspect of the experiment has the most influence on the overall uncertainty in L_M ?

.....

- How could this aspect be modified to improve the precision and/or accuracy of the final result?

.....

.....

- What other improvements can you suggest?

.....

.....

- Very briefly: What other methods can be used to determine L_M ?
How do the results of such experiments compare with yours?

.....

.....

OBJECTIVES FOR UNIT 4

After you have worked through this Unit, you should be able to:

- 1 Explain the meaning of, and use correctly, all the terms flagged in the text.
- 2 Distinguish between random and systematic uncertainties, and between precise and accurate measurements.
- 3 Make reasonable estimates of the uncertainties in an experiment and decide which uncertainties are likely to be the dominant ones.
- 4 Present the results of an experiment in the form of Tables and/or graphs, and identify errors or deficiencies of presentation in respect of:
 - (a) data—the recording of data in tabular form, the estimation of uncertainties, the number of significant figures quoted;
 - (b) graphs—the choice of variables, axes and scales, the manner in which the axes are labelled, the points and error bars plotted and the appropriate curve(s) drawn.
 - 5 Understand the meaning of the term ‘standard deviation’ (the formula need *not* be recalled).
 - 6 Write a report of practical work, containing
 - (a) an abstract
 - (b) a concise description of experimental procedure
 - (c) tabular and/or graphical presentation of data, as appropriate
 - (d) estimates of uncertainties in individual measurements
 - (e) a final result, with associated confidence limit, expressed to an appropriate number of significant figures
 - (f) a conclusion

ITQ ANSWERS AND COMMENTS

ITQ 1 You should have remembered from Units 2 and 3 that entries in Tables are normally given as dimensionless numbers. (The same is true of points plotted on graphs.) Column headings should therefore give the quantity being tabulated, divided by the units and an appropriate power of ten.

In this case a suitable layout would be:

Planet	Mean distance from Sun/ 10^{11} m	Orbital period/Earth years
Mercury	0.58	0.24
Venus	1.08	0.62
Earth	1.50	1.00

Comments (i) The scale in the distance column has been chosen to keep as many of the entries as possible between 1 and 10, and this is a good rule of thumb. However, it would not be *wrong* to head this column ‘mean distance from Sun/ 10^{12} m’, and to record the readings as 0.058, 0.108 and 0.150.

(ii) In the question the measurements for the orbital periods of Mercury and Venus were given to two decimal places. In the Table above, the orbital period of Earth has also been quoted to this degree of precision, even though the value for the period of the Earth (unity) is *exact*, since the orbital periods are given in Earth years (we would, therefore, be justified in putting an infinite number of zeros after the decimal point). We will return to the subject of how to quote measurements in Section 4.

ITQ 2 The graph of Figure 7 begs at least one question: ‘Why does the graph become non-linear for masses greater than about 35 kg?’ You might, therefore, want to see whether you could discover any physical differences in the response of the wire to masses over 35 kg and masses under 35 kg, in addition to the difference in extension. (In fact, you would find that for extensions corresponding to the linear portion of the

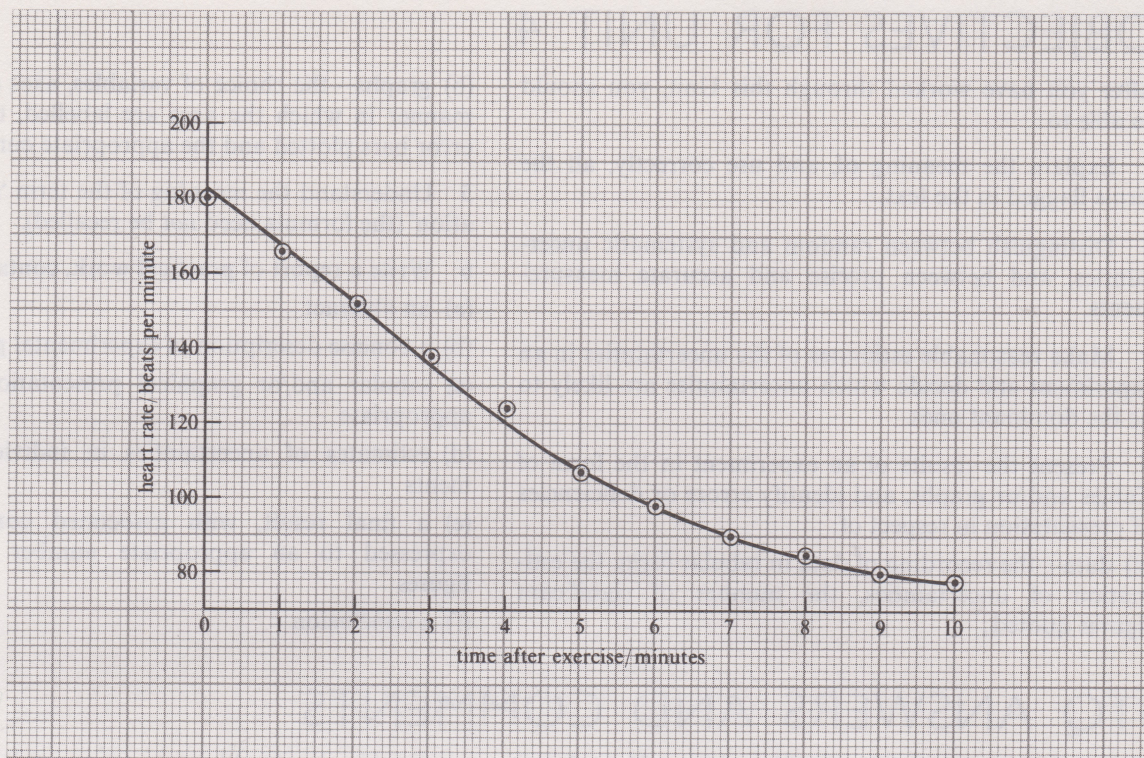
graph, the wire would return to its original length when the load was removed. Much larger loads, however, would cause permanent elongation of the wire.) In a more extensive set of experiments you might want to find out how the extension varied with the original length, shape and cross-sectional area of the wire, or with the purity of the copper. You might also want to consider the influence of external factors, such as temperature, on the results.

- ITQ 3** (a) $1 \text{ km in } 50 \text{ s} = 0.02 \text{ km s}^{-1}$
 (b) 0.018 km s^{-1}
 (c) $3 \times 10^5 \text{ km s}^{-1}$

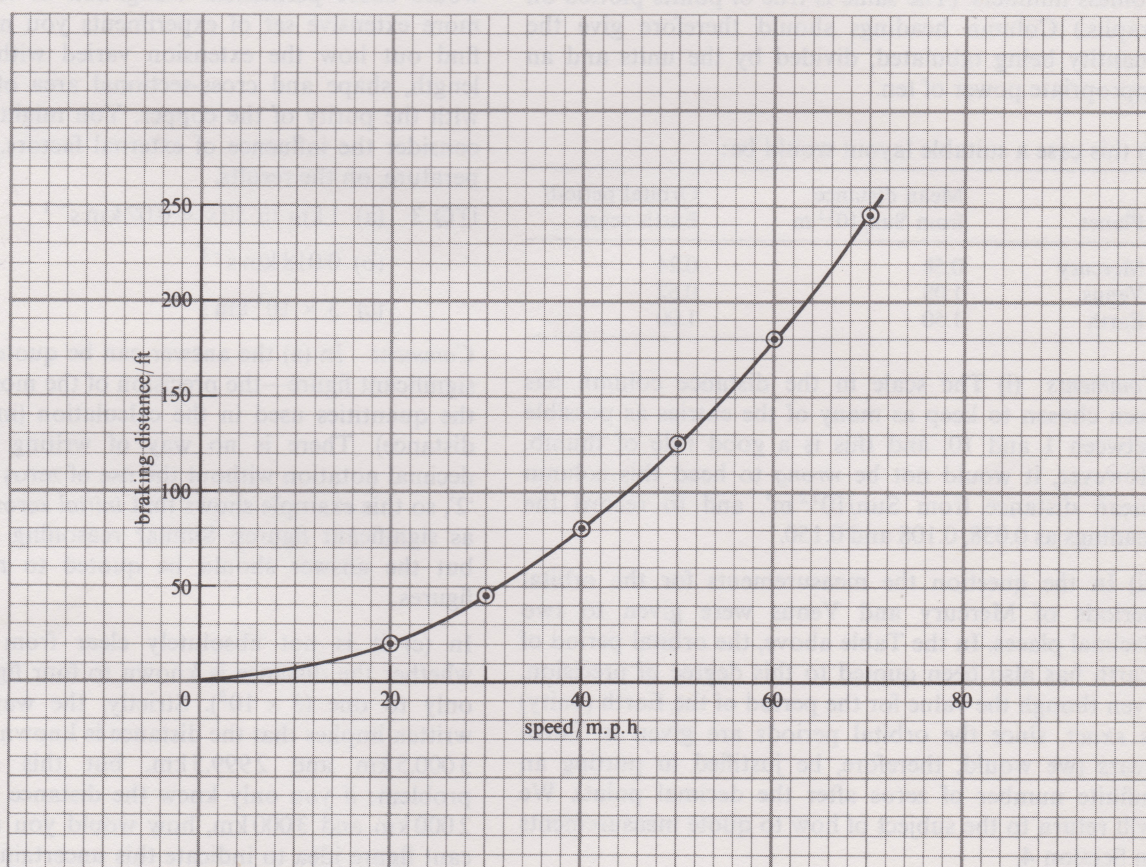
Comment In (a) the answer can be quoted to only one significant figure—the precision of the most uncertain of the quantities used in the calculation (in this case the distance). There is no way of writing this speed in decimal notation without the use of zeros in front of the ‘2’, so this example shows that *initial* zeros do not count as significant figures. Similar reasoning applies to (b), but the answer should be quoted to *two* significant figures.

In (c) it is not absolutely clear from the question whether the distance is known to four figures (3000) or only to one (3×10^3). Strictly, the way it has been written implies that the distance is known to lie between 3000.5 km and 2999.5 km, but this still leaves a problem: if you only knew the distance to be between 2000 km and 3000 km, how would you use the significant figure idea to indicate this uncertainty? *You would not have had this dilemma if scientific notation had been used: 3×10^3 km corresponds to one significant figure, and 3.000×10^3 km to four significant figures.*

Either way, the answer to (c) is governed by the fact that the time (0.01 s) is quoted to just one significant figure, so the speed cannot be given more accurately than that: $3 \times 10^5 \text{ km s}^{-1}$ is the only unambiguous way to write this result.



(a)



(b)

FIGURE 27 Answer to SAQ 1.

SAQ ANSWERS AND COMMENTS

SAQ 1 (a) A suitable choice of scale is shown in Figure 27a. Here, it is not necessary to scale the vertical axis right down to zero (heart rate does not fall to zero!). The lowest heart rate measured in the experiment was 78 beats per minute, so it is sensible to start the vertical scale at about 70 beats per minute. This allows some space for the curve to go below the lowest data point if the scatter of the points requires this.

(b) A suitable choice of scales is shown in Figure 27b. In this case it is sensible to include the origin, since common sense tells us that the curve should go through it (a car that isn't moving will not require any time to stop). By interpolation, the braking distance at a speed of 45 m.p.h. is 100 ft.

SAQ 2 Six key faults in Figure 16a are:

- 1 the axes are wrongly chosen—the independent variable is time and it should be plotted on the horizontal axis;
- 2 the horizontal axis has no label;
- 3 the vertical axis should have the label 'time/s';
- 4 the choice of scale is poor on the horizontal axis;
- 5 several points are incorrectly plotted;
- 6 the points should be crosses or small circled dots, not large blobs. A sensible graphical presentation of the data is shown in Figure 28.

The speed of the car after 25 s is 74.6 km h^{-1}

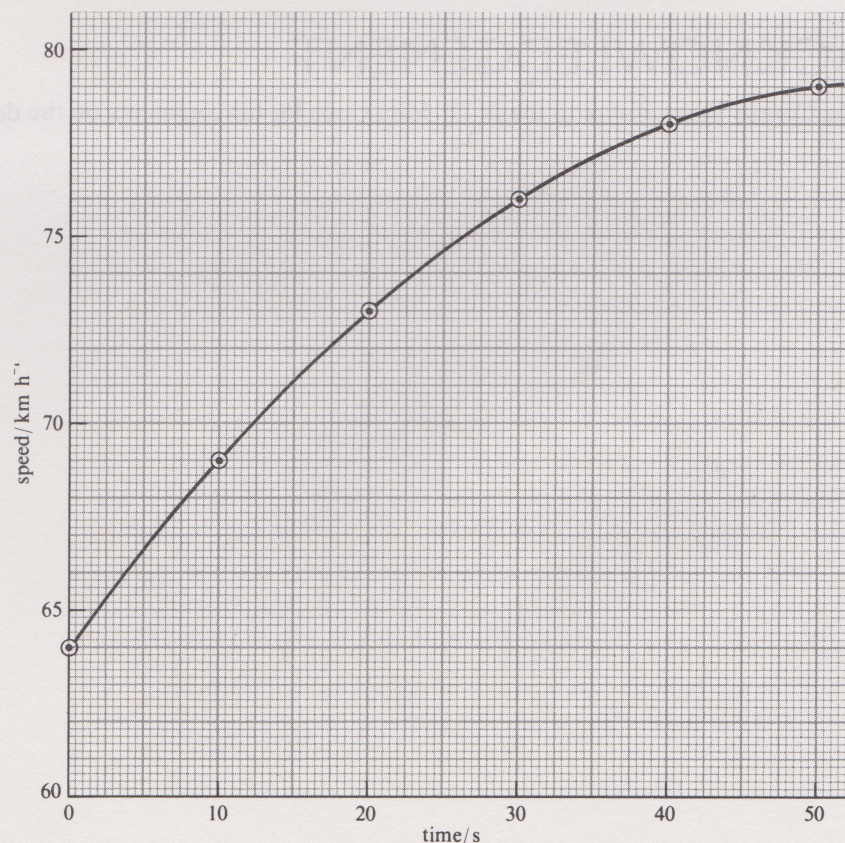


FIGURE 28 Answer to SAQ 2.

INDEX FOR UNIT 4

- abstract of a report, 29
- accuracy and precision, 17–20
- accurate measurement, 18, 19**
- average, calculating an, 19, 26
- axes of graph, 11, 13–15, 24
- calibration, 19**
- checking calculations, 25
- control group, 16**
- copper wire, stretching of, 9–12
- data,
 - handling, 9–23
 - quoting experimental results, 21
 - standard deviation, 25–7
 - tabulating, 8, 9
 - see also graphs; practical work; uncertainties
- dependent variable, 11, 24**
- designing experiments, 4–6
- error bars, 22, 23, 24**
- errors, systematic, 18–19
 - see also uncertainties
- estimating uncertainties, 19–20
- experiments, doing; see practical work
- extrapolation, 11
- field notes, 7–8
- graphs, 9–15, 22–4
 - plotting, 9–15, 24
- graphical representation of
 - uncertainties, 22–3
- scale, choosing a, 13, 14–15, 24
- identifying uncertainties, 16–17
- independent variable, 11, 24**
- instrumental limitations in
 - experiments, 17
- intercept, 12, 13**
- interpolation, 11**
- manipulating statistics, 13
- measurement, in practical work, 4, 17–20
- micrometer, 5
- Moon, measuring distance to, 7, 27–34
- observation, 4
- parallax errors, 17
- planning experiments, 6–7, 28
- practical work and experiments
 - designing, 4–6
 - measurement, 4, 17–20
 - observation, 4
 - planning, 6–7, 28
 - records of, 7–10
 - rehearsing, 7
 - reports, writing, 27–34
 - see also data
- precise measurement, 17, 18, 20**
- qualitative work, 4
- quantitative result, 4
- quoting results, 21
- random uncertainties, 18, 19–20
- records, 7–8
- rehearsal for experiment, 7
- reports, 27–34
- results
 - quoting, 21
 - reporting, 27–34
- sampling, 4–6, 16
- scale of graph, choosing, 13, 14–15, 24
- significant figures, 21**
- standard deviation, 25, 26–7**
- statistical significance tests, 17
- statistics, manipulation of, 13
- systematic errors, 18–19
- Tables of data, 8, 9
- uncertainties, 18
 - estimating, 19–20
 - graphical representation of, 22–3
 - identifying, 16–17
 - limits, 22, 23
 - quoting, 21
 - random, 18, 19–20
 - see also systematic error
- variable,
 - dependent, 11, 24
 - independent, 11, 24
- Vernier scale, 5
- volumes, measuring, 6
- written reports and records, 7–8, 27–34

ACKNOWLEDGEMENT

The Course Team wishes to thank Dr John Cope for his comments on the drafts of Units 1 to 4.